Validation of satellite sounder environmental data records: Application to the Cross-track Infrared Microwave Sounder Suite

Nicholas R. Nalli,[1,2] Christopher D. Barnet,[3] Anthony Reale,[4] David Tobin,[5] Antonia Gambacorta,[1,2] Eric S. Maddy,[3,2] Everette Joseph,[6] Bomin Sun,[1,4] Lori Borg,[5] Andrew K. Mollner,[7] Vernon R. Morris,[6] Xu Liu,[8] Murty Divakarla,[1,2] Peter J. Minnett,[9] Robert O. Knuteson,[5] Thomas S. King,[1,2] and Walter W. Wolf[2]

Received 27 June 2013; revised 30 October 2013; accepted 28 November 2013; published 26 December 2013.

[1] The Joint Polar Satellite System (JPSS) Cross-track Infrared Microwave Sounder Suite (CrIMSS) is an advanced operational satellite sounding system concept comprised of the Cross-track Infrared Sounder and the Advanced Technology Microwave Sounder. These are synergistically designed to retrieve key environmental data records (EDR), namely atmospheric vertical temperature, moisture, and pressure profiles. CrIMSS will serve as the low-Earth orbit satellite sounding system, starting with the Suomi National Polar-orbiting Partnership (S-NPP) satellite and spanning the JPSS-1 and JPSS-2 satellites. This paper organizes the general paradigm for validation of atmospheric profile EDR retrieved from satellite nadir sounder systems (e.g., CrIMSS) as a synthesis of complementary methods and statistical assessment metrics. The validation methodology is ordered hierarchically to include global numerical model comparisons, satellite EDR intercomparisons, radiosonde matchup assessments (conventional, dedicated, and reference), and intensive campaign "dissections." We develop and recommend the proper approach for computing profile statistics relative to correlative data derived from high-resolution in situ data (viz., radiosondes) reduced to forward model layers. The standard statistical metrics used for EDR product assessments on "coarse layers" are defined along with an overview of water vapor weighting schemes and the use of averaging kernels. We then overview the application of the methodology to CrIMSS within the context of the JPSS calibration/validation program, with focus given to summarizing the core data sets to be used for validation of S-NPP sounder EDR products.

Citation: Nalli, N. R., et al. (2013), Validation of satellite sounder environmental data records: Application to the Cross-track Infrared Microwave Sounder Suite, *J. Geophys. Res. Atmos.*, *118*, 13,628–13,643, doi:10.1002/2013JD020436.

1. Introduction

[2] The Joint Polar Satellite System (JPSS) is a U.S. National Oceanic and Atmospheric Administration (NOAA) operational satellite mission in collaboration with joint international partnerships and the U.S. National Aeronautics and Space Administration (NASA). The primary mission of the JPSS program is to support NOAA's weather and climate mission by providing operational environmental data records (EDR) (directly or indirectly) to NOAA data users. Advanced sensors to be flown on JPSS low-Earth orbit environmental satellites include the Cross-track Infrared Microwave Sounder Suite (CrIMSS), a passive nadir sounding system [e.g., *Smith et al.*, 2009] consisting of the Cross-track Infrared Sounder (CrIS), and the Advanced Technology Microwave Sounder (ATMS). The CrIMSS sounding system is designed to measure well-calibrated infrared (IR) and microwave (MW) radiances (called sensor data records, SDRs) for synergistically retrieving atmospheric vertical profile EDR under nonprecipitating conditions (clear, partly cloudy, and cloudy) in much the same manner as earlier sounders, namely the MetOp-A and -B Infrared

[1]IMSG, Inc., Rockville, Maryland, USA.
[2]NOAA/NESDIS/STAR, College Park, Maryland, USA.
[3]STC, Columbia, Maryland, USA.
[4]NOAA/NESDIS/STAR, Suitland, Maryland, USA.
[5]University of Wisconsin-Madison, Madison, Wisconsin, USA.
[6]Howard University, Washington, Dist. of Columbia, USA.
[7]The Aerospace Corporation, El Segundo, California, USA.
[8]NASA Langley Research Center, Hampton, Virginia, USA.
[9]RSMAS, University of Miami, Miami, Florida, USA.

Corresponding author: N. R. Nalli, NOAA Center for Weather and Climate Prediction, IMSG, Inc., 5830 University Research Court, College Park, MD 20740-3818, USA. (Nick.Nalli@noaa.gov)

2169-897X/13/10.1002/2013JD020436

Table 1. JPSS Specification Performance Requirements[a] for AVTP

AVTP Measurement Uncertainty (3σ) Layer Average Temperature Error	
Boundaries	Thresholds
"Partly Cloudy" (IR + MW)	
Surface to 300 hPa	1.6 K / 1 km layer
300 hPa to 30 hPa	1.5 K / 3 km layer
30 hPa to 1 hPa	1.5 K / 5 km layer
1 hPa to 0.5 hPa	3.5 K / 5 km layer
"Cloudy" (MW only)	
Surface to 700 hPa	2.5 K / 1 km layer
700 hPa to 300 hPa	1.5 K / 1 km layer
300 hPa to 30 hPa	1.5 K / 3 km layer
30 hPa to 1 hPa	1.5 K / 5 km layer
1 hPa to 0.5 hPa	3.5 K / 5 km layer

[a]Source: Joint Polar Satellite System (JPSS) program level 1 requirements document—preliminary, version 4.7, 22 September 2011, NOAA/NESDIS, NASA.

Atmospheric Sounding Interferometer (IASI) [*Cayla*, 1993; *Maddy et al.*, 2011; *Hilton et al.*, 2012] and the EOS-Aqua Atmospheric Infrared Sounder (AIRS) [*Aumann et al.*, 2003; *Chahine et al.*, 2006]. The primary CrIMSS EDR are atmospheric vertical temperature, moisture, and pressure profiles (AVTP, AVMP and AVPP, respectively). The CrIS instrument is an advanced Fourier transform spectrometer that measures high-resolution IR spectra in 1305 channels over three bands spanning $\nu \simeq [650, 2550]$ cm^{-1} (high spectral resolution is hereafter simply referred to as "hyperspectral"). The ATMS is a MW sounder with 22 channels ranging from 23 to 183 GHz [*Weng et al.*, 2012]. These two instruments operate in a "golf ball" formation analogous to AIRS, with ATMS fields of view (FOVs) resampled to match the location and size of the 3×3 CrIS FOVs for retrievals under partly cloudy conditions [e.g., *Susskind et al.*, 2003]. There have been two CrIMSS retrieval algorithms independently developed for operational use (discussed more in Appendix A). The first algorithm is the NOAA Unique CrIS/ATMS Processing System (NUCAPS) developed at NOAA/NESDIS/STAR (e.g., A. Gambacorta et al., The NOAA Unique CrIS/ATMS processing System (NUCAPS): Algorithm description and preliminary validation, *Journal of Atmospheric and Oceanic Technology*, manuscript in preparation, 2013). The second is an optimal estimation (OE) [*Rodgers*, 1976] physical retrieval algorithm [*Lynch et al.*, 2009] originally developed by Atmospheric and Environmental Research, Inc. (AER) [*Moncet et al.*, 2005] for the CrIS sensor contractor, which was subsequently enhanced and initially integrated into the Interface Data Processing Segment (IDPS) by the prime contractor Northrop Grumman Aerospace Systems (NGAS) (hereafter "original IDPS" or "original CrIMSS" algorithm).

[3] Users of environmental satellite data products have changed over the past decade since the launch of these advanced hyperspectral sounders. The needs of the traditional user base have shifted, and new users have emerged. While it is understood that environmental satellite sensors measure spectral radiances (SDRs), the primary objective nonetheless is to observe and predict the geophysical state of the Earth system. Radiances assimilated within numerical weather prediction (NWP) models are used for one-dimensional, three-dimensional, or four-dimensional variational analysis. Beyond this, some users prefer to utilize directly the retrieved Level 2 EDR products for observations [*Goldberg et al.*, 2003]. For example, weather forecast offices utilize EDR via the Advanced Weather Interactive Processing System, and EDR data sets are used for numerous environmental research studies [e.g., *Pagano*, 2013]. Furthermore, EDR retrievals based upon cloud-cleared radiances (CCRs) [e.g., *Susskind et al.*, 2003; *Maddy et al.*, 2011] have been shown to have a positive impact on forecast models [e.g., *Le Marshall et al.*, 2008; *Reale et al.*, 2008b, 2009, 2012a; *Zhou et al.*, 2010].

[4] *Fetzer et al.* [2003] describe validation as "the process of ascribing uncertainties to these radiances and retrieved quantities through comparison with correlative observations." Because the EDR retrieval (inverse problem) and radiative transfer calculation (forward problem) are essentially transformations between radiance space and geophysical space, EDR validation can constitute an implicit validation of the SDR and forward radiative transfer model (RTM). EDR validation also facilitates validation of CCRs and cloud detection algorithms, for example, by using subsets of scenes that are cloud free or cloud masked [e.g., *Divakarla et al.*, 2006]; by comparing CCRs against forward calculations (i.e., calc – obs) [e.g., *Nalli et al.*, 2013]; by analyzing the behavior of cloud-parameter retrievals [e.g., *Nalli et al.*, 2013]; or by analyzing surface temperature EDR and high-resolution imager data [e.g., *Maddy et al.*, 2011]. For the above reasons, EDR validation has relevance to all users of environmental satellite data.

[5] To accommodate a phased schedule and ensure that the CrIMSS EDR would comply with the mission Level 1 requirements (Tables 1 and 2), a calibration/validation (cal/val) plan for the JPSS Suomi National Polar-orbiting Partnership (S-NPP) satellite was devised [*Barnet*, 2009], leveraging upon previous sounder cal/val work [e.g., *Fetzer et al.*, 2003; *Hagan and Minnett*, 2003; *Hagan et al.*, 2004; *Fetzer*, 2006; *Tobin et al.*, 2006a, 2006b; *Nalli et al*, 2006; *Divakarla et al.*, 2006; *Taylor et al.*, 2008; *Seidel et al.*, 2009; *Nalli et al.*, 2011; *Maddy et al.*, 2012; *Newman et al.*, 2012; *Reale et al.*, 2012b], along with earlier plans drafted by NGAS and workshops initiated by the former Integrated Program Office (IPO) [e.g., *Mango et al.*, 2001]. In this paper it is our intention to formalize this basic framework incorporating the various sounder

Table 2. JPSS Specification Performance Requirements[a] for AVMP

AVMP Measurement Uncertainty (3σ) The 2 km Layer Average Mixing Ratio Proportional Error	
Boundaries	Thresholds
"Partly Cloudy" (IR + MW)	
Surface to 600 hPa	greater of 20% or 0.2 g kg^{-1}/2 km layer
600 hPa to 300 hPa	greater of 35% or 0.1 g kg^{-1}/2 km layer
300 hPa to 100 hPa	greater of 35% or 0.1 g kg^{-1}/2 km layer
"Cloudy" (MW only)	
Surface to 600 hPa	greater of 20% or 0.2 g kg^{-1}/2 km layer
600 hPa to 400 hPa	greater of 40% or 0.1 g kg^{-1}/2 km layer
400 hPa to 100 hPa	greater of 40% or 0.1 g kg^{-1}/2 km layer

[a]Source: Joint Polar Satellite System (JPSS) program level 1 requirements document—preliminary, version 4.7, 22 September 2011, NOAA/NESDIS, NASA.

validation techniques as a general methodology and standard reference for validation of EDR from passive sounder systems. The validation methodology and data sets are organized and described in section 2, followed in section 3 by a treatise on the statistical techniques used for EDR profile validation, including the proper treatment of correlative data and sensor averaging kernels. Plans for the specific application of the validation methodology toward validation of the S-NPP operational EDR (AVTP, AVMP, and Ozone) is then discussed in section 4, with focus given to the ongoing acquisition of validation data sets. Details on the retrospective application of the methodology toward the Beta and Provisional Maturity status of the original IDPS S-NPP CrIMSS operational EDR products can be found in the related EDR paper by *Divakarla et al*. [2014] in this Special Section.

2. Validation Methodology

[6] This section organizes what may be considered the standard paradigm for IR sounder validation. In the subsections that follow, well-established validation techniques and data sets are arranged in an order of increasing maturity in the EDR product validation process in what may be considered a validation "hierarchy." Techniques at the beginning of the hierarchy are those typically utilized at the beginning stages of EDR validation (i.e., early in the satellite mission), whereas those near the end are typically applied in the later stages.

2.1. Numerical Model Global Comparisons

[7] Matchups of EDR with numerical weather prediction (NWP) model output (analysis and forecast interpolated to EDR) enable the rapid acquisition of large, global data sets during "Focus Days" (i.e., days selected for the archival of global SDRs that are used for reprocessing of EDR using the latest versions of nonoperational "offline" code) and as such are extremely useful for early evaluation of the algorithm and identifying gross problem areas [e.g., *Fetzer et al.*, 2003]. NWP models include the European Centre for Medium-Range Weather Forecasts (ECMWF) and National Center for Environmental Prediction Global Forecasting System. Such analyses are useful in identifying regional or spectral biases. However, NWP models do not constitute independent correlative data. First, the models themselves assimilate satellite radiances, including CrIS/ATMS. Second, NWP models have been used both for training regression algorithms [e.g., *Goldberg et al.*, 2003] (both linear and nonlinear) and in radiance bias corrections used for forward model tuning (A. Gambacorta et al., A methodology for computing systematic biases of the top of atmosphere brightness temperature calculations Part I: Infrared brightness temperature computations, manuscript in preparation JGR).

2.2. Satellite EDR Intercomparisons

[8] Matchups can also be obtained from the EDR of other sensors or algorithms. Like NWP model comparisons, this approach also allows the acquisition of large, global data samples, although more time may be necessary, depending on orbital overlaps. Thus, these can also facilitate early consistency checks but depending on the sensor/algorithm can provide a more independent data set as well as additional confidence in the case of previously validated EDR products. Furthermore, such intercomparisons can also highlight the merits of different retrieval sensors/algorithms, including the benefits of hyperspectral IR. Example sounder EDR data sets include AIRS, IASI, and the Advanced TIROS Operational Vertical Sounder [e.g., *Li et al.*, 2000; *Reale et al.*, 2008a]. A critical limitation of these data for validation, however, is that they inherently possess similar retrieval error characteristics [*Fetzer et al.*, 2003] and otherwise require proper rigorous treatment of each sensor's averaging kernels [*Rodgers and Connor*, 2003]. Another source of satellite EDR for intercomparison are those from Global Positioning System (GPS) radio occultation (RO), for example, those obtained from the Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC) [e.g., *Anthes et al.*, 2008; *Sun et al.*, 2010, 2013; *Wang et al.*, 2013], which constitute a fundamentally different measurement from nadir sounders, having very good temperature accuracy for the upper troposphere/lower stratosphere. Drawbacks of using GPS RO are the limitation to altitudes higher than 300 hPa and that global sampling is nonuniform with fewer samples in the tropical regions.

2.3. Conventional RAOB Matchup Assessments

[9] Balloon-borne radiosonde observations (RAOBs) launched twice per day at World Meteorological Organization upper air observing stations are the conventional method for obtaining atmospheric soundings on synoptic to global scales. Because some of the RAOBs inevitably coincide with satellite overpasses, conventional RAOB matchups are a source of validation data that supercede NWP model comparisons for providing global zonal assessments and long-term characterization [*Divakarla et al.*, 2006; *Reale et al.*, 2012b]. However, it can take several months of RAOB matchup collocations to accumulate an adequate sample size. Typically, $\simeq$1000 RAOBs can be matched up with overpasses within $\pm$6 h and 250 km radius per day [*Reale et al.*, 2012b]. However, for any given orbit, these tend to have a particular regional sampling and asymmetric coverage [e.g., *Fetzer et al.*, 2003; *Divakarla et al.*, 2006], along with a skewed distribution toward Northern Hemisphere continental locations [e.g., *Divakarla et al.*, 2006; *Reale et al.*, 2012b]. Assessments based on conventional RAOB matchups will also suffer from (1) significant mismatch/noncoincidence errors, assuming the loose space-time matchup criteria mentioned above that are typically employed [*Fetzer et al.*, 2003; *Pougatchev*, 2008; *Sun et al.*, 2010], (2) limited vertical resolution in the available data (typically mandatory and significant levels), and (3) a nonuniform set of radiosonde types used in the data sample [e.g., *Divakarla et al.*, 2006; *Miloshevich et al.*, 2006; *Sun et al.*, 2010, 2013], including those with error characteristics that are inadequately analyzed and documented or otherwise have suboptimal accuracy [*Miloshevich et al.*, 2006; *Sun et al.*, 2010]. For these reasons, it is important to supplement conventional RAOB matchups with dedicated and/or reference RAOBs as discussed in the next subsection.

2.4. Dedicated and Reference RAOB Matchup Assessments

[10] Like their name implies, dedicated RAOBs are derived from radiosondes dedicated to satellite overpasses for the explicit purpose of validation [e.g., *Fetzer et al.*, 2003; *Tobin et al.*, 2006a; *Nalli et al.*, 2011]. These consist typically of state-of-the-art Vaisala GPS rawinsondes, currently RS92-SGP, that are launched to coincide with satellite overpasses, usually $\simeq$15–60 min in advance to allow for balloon ascent through the lower troposphere [e.g., *Tobin et al*, 2006a; *Nalli et al.*, 2011]. Dedicated RAOBs thus contain only minimal mismatch errors [*Fetzer et al.*, 2003], and they also include GPS-derived altitudes to allow validation of AVPP EDR as a function of geometric altitude. RS92 radiosondes also have well-known error characteristics and optimal accuracy among the available types [*Miloshevich et al.*, 2006]. Ideally, dedicated RAOBs may also include atmospheric state "best estimates" [*Tobin et al.*, 2006a] or "merged" soundings (sounding profiles obtained through a combination of RAOB and some other source such as lidar or numerical model), as well as "reference radiosondes" such as cryogenic frostpoint hygrometers (CFH) [*Seidel et al*, 2009] or the Vaisala RR01 radiosonde currently under development [*Turtiainen et al.*, 2010]. Vaisala RS92 sondes may also qualify as reference sondes when corrections are applied for calibration and time lag errors [*Miloshevich et al*, 2006].

[11] For these reasons, dedicated RAOBs can provide detailed performance assessment and allow spot checks of global model or conventional RAOB analyses and hence provide a "bootstrap" verification of the global model analyses described in section 2.1. Obvious limitations of dedicated RAOBs include small sample sizes and spatiotemporal coverage [*Fetzer et al.*, 2003]. However, when combined with numerical model and conventional RAOB analyses, dedicated RAOBs nevertheless provide independent verification that assessments are valid (i.e., statistics versus longitude, land/ocean, day/night, versus view angle, and such) by confirming those results under representative conditions with smaller sample sizes. Dedicated RAOBs may thus be thought of as a "final exam" for accuracy assessment. Example-dedicated RAOB sites include U.S. Department of Energy (DOE) Atmospheric Radiation Measurement (ARM) sites [*Tobin et al.*, 2006a; *Long et al.*, 2013], sites of opportunity (see section 4.2.3), and ideally Global Climate Observing System Reference Upper Air Network (GRUAN) [*Seidel et al.*, 2009]. Dedicated RAOB data currently being acquired for CrIMSS validation, including best estimate and merged soundings, are discussed more in section 4.2.3.

2.5. Intensive Field Campaign Dissections

[12] Well-designed intensive validation field campaigns constitute the "top" of the validation hierarchy because they can provide the most ideal data set for sounder EDR and SDR validation, namely a set of complementary data that taken together can provide the most thorough specification of the atmospheric state. Intensive campaigns minimally include dedicated RAOBs, especially those not assimilated into NWP models (see section 4.2.4) but also typically include any number of ancillary data sets such as ozonesondes [e.g., *Adam et al.*, 2010; *Nalli et al.*, 2011], Marine Atmospheric Emitted Radiance Interferometers (M-AERI) [e.g., *Minnett et al.*, 2001; *Nalli et al.*, 2006], lidars [e.g., *Adam et al.*, 2010; *Mollner et al.*, 2013], sunphotometers [e.g., *Nalli et al.*, 2006, 2011], microwave radiometers (MWR), and others. Intensive campaigns for sounder validation also ideally include underflights of an aircraft IR spectrometer sounder such the National Polar-orbiting Operational Environmental Satellite System Atmospheric Sounder Testbed-Interferometer (NAST-I) and Scanning High-resolution Interferometer Sounder (S-HIS), which allow for direct SDR cal/val [e.g., *Tobin et al.*, 2006b]. Field campaigns that include data sets such as those mentioned can allow for detailed performance specification and EDR validation "dissections" (i.e., dissection of the retrieval to separate and identify the specific contributors to the cumulative retrieval error). Examples of intensive validation field campaigns are found in *Whiteman et al.* [2006], *Taylor et al.* [2008], *Adam et al.* [2010], *Newman et al.* [2012], and *Nalli et al.* [2006, 2011]. A somewhat new paradigm for intensive validation campaigns that has become more prominent are scientific "campaigns of opportunity" (e.g., see section 4.2.4). Campaigns of opportunity constitute an improvised approach by leveraging collaborative science partnerships beyond simply that of validation in an effort to distribute cost and risk. Value-added benefits to this approach include science synergism, engagement of the broader science community, and new cross-disciplinary opportunities involving the application of operational satellite sounder EDR, although a major drawback is reduced control for applying any optimal sampling strategies that are oriented toward validation.

3. Assessment Methodology

[13] Given validation data sets as described in section 2 above, in this section we lay out the basis for a proper computation of validation statistics of retrieved profile EDR relative to a correlative measurement (e.g., RAOBs). The approach overviewed here has been developed and refined based upon AIRS and IASI sounder science team validation experience.

3.1. Statistical Metrics

[14] A proper "apples-to-apples" comparison between high-resolution "truth" measurements (e.g., RAOBs) with a satellite-retrieved state (i.e., EDR) requires that the measurement first be reduced to forward model effective-layer quantities, the details of which are given in Appendix B. Given profiles on forward model layers (i.e., the correlative true state) matched up with EDR retrievals on the same layers, error statistics are then computed as detailed below.

[15] Level 1 AVTP and AVMP accuracy requirements are defined over what are referred to as "coarse layers," roughly 1–5 km layers for tropospheric AVTP (Table 1) and 2 km layers for AVMP (Table 2). Note that the coarse layers specification roughly corresponds to the effective vertical resolution of hyperspectral sounders (discussed more in section 3.2). Because temperature is measured on an absolute scale over which the relative precision in measurement is uniform, temperature (AVTP) statistics on coarse layers are relatively straightforward. However, the relative precision of water vapor measurement depends greatly upon both the range of absolute water vapor content and, in the case of

RAOBs, the ambient temperature [*Miloshevich et al.*, 2006], and thus, the calculation of AVMP statistics have assumed different expressions, depending on goals and rationale.

3.1.1. Temperature

[16] The root-mean-square (RMS) temperature difference provides a single metric that includes all deviations (systematic and random) in the matchup sample. Forward model effective-layer temperatures T_L are averaged to obtain coarse layer values

$$T_{\mathfrak{L}} = \frac{\sum_{L(\mathfrak{L})} \ln(P_L/P_{L-1})\, T_L}{\ln(P_{\mathfrak{L}}/P_{\mathfrak{L}-1})}\,, \tag{1}$$

where $L(\mathfrak{L})$ are the forward model effective layers that fall within the coarse layer $\mathfrak{L}$. Defining the deviation of a single AVTP retrieval from a correlative profile on $m_{\mathfrak{L}} \simeq 20$ coarse layers at a matchup location j as

$$\Delta T_{\mathfrak{L},j} \equiv \widehat{T}_{\mathfrak{L},j} - T_{\mathfrak{L},j}\,, \quad \mathfrak{L} = 1, 2, \ldots, m_{\mathfrak{L}}\,, \tag{2}$$

where retrieved quantities are denoted with a hat, the RMS deviation is given by

$$\mathrm{RMS}(\Delta T_{\mathfrak{L}}) = \sqrt{\frac{1}{n_j}\sum\nolimits_{j=1}^{n_j} (\Delta T_{\mathfrak{L},j})^2}\,, \tag{3}$$

where n_j is the matchup sample size. The measure of sample central tendency is given by the bias statistic, which is calculated simply as the mean difference

$$\mathrm{BIAS}(\Delta T_{\mathfrak{L}}) \equiv \overline{\Delta T}_{\mathfrak{L}} = \frac{1}{n_j}\sum_{j=1}^{n_j} \Delta T_{\mathfrak{L},j}\,. \tag{4}$$

Variability about the mean is measured by the standard deviation, σ, which can be derived from the RMS and BIAS statistics in equations (3) and (4) as

$$\mathrm{STD}(\Delta T_{\mathfrak{L}}) \equiv \sigma(\Delta T_{\mathfrak{L}}) = \sqrt{[\mathrm{RMS}(\Delta T_{\mathfrak{L}})]^2 - [\mathrm{BIAS}(\Delta T_{\mathfrak{L}})]^2}\,. \tag{5}$$

3.1.2. Water Vapor and Ozone

[17] As pointed out by *Maddy and Barnet* [2008], IR sensitivity to a trace gas species x is more linear for relative changes ($\delta x/x$) as opposed to absolute changes (δx), this being a consequence of the exponential dependence of extinction on layer abundance according to Beer's law. Thus, the skill in AVMP, ozone, and other trace gas profile retrievals are also usually defined in a relative sense. Because the analysis of relative changes tends to inflate the impact of atmospheres with small amounts, it has been desirable to devise weighting schemes that reduce the skewing of the sample statistics toward atmospheres with very small amounts of H_2O (or O_3).

[18] For computing statistics on the coarse layers, H_2O and O_3 profiles (both retrieval and correlative) are first converted to the forward model layer abundances (molecules/cm^2), which are then summed over a coarse layer to yield coarse layer abundances. For H_2O, we denote forward model layer molecular abundances as $\overline{N}_{w,L}\,\delta z_L$, where $\overline{N}_{w,L}$ is the layer mean number density and δz_L is the layer thickness. The coarse layer mass abundances $q_{\mathfrak{L}}$ (in g/cm^2) are then calculated as

$$q_{\mathfrak{L}} = \frac{M_w}{N_A}\sum_{L(\mathfrak{L})} \overline{N}_{w,L}\,\delta z_L\,, \tag{6}$$

where M_w is H_2O molecular mass, N_A is Avogadro's number, and $L(\mathfrak{L})$ is as above in (1). The bottom layer is conventionally multiplied by the factor

$$f_{l_b} \equiv \frac{p_s - P_{l_b-1}}{P_{l_b} - P_{l_b-1}}\,, \tag{7}$$

where the index l_b denotes the bottom level (as defined in Appendix B).

[19] Given coarse layer abundances, $q_{\mathfrak{L}}$, the fractional deviation is taken to be the absolute deviation divided by the observed (e.g., RAOB) value

$$\Delta q_{\mathfrak{L},j} \equiv \frac{\hat{q}_{\mathfrak{L},j} - q_{\mathfrak{L},j}}{q_{\mathfrak{L},j}}\,, \quad \mathfrak{L} = 1, 2, \ldots, m_{\mathfrak{L}}\,. \tag{8}$$

From here, formulas for RMS, BIAS, and STD statistics can be obtained simply by substituting $\Delta q_{\mathfrak{L}}$ for $\Delta T_{\mathfrak{L}}$ in (3), (4), and (5). However, the denominator in equation (8) can result in large $\Delta q_{\mathfrak{L},j}$ in dry atmospheres (e.g., middle to upper troposphere or polar regions) and thereby skew the statistics toward these cases. Therefore, sounder science team cal/val convention has been to implement more general weighted means as the basis for the statistics (rather than arithmetic means):

$$\mathrm{RMS}(\Delta q_{\mathfrak{L}}) = \sqrt{\frac{\sum_{j=1}^{n_j} W_{\mathfrak{L},j}(\Delta q_{\mathfrak{L},j})^2}{\sum_{j=1}^{n_j} W_{\mathfrak{L},j}}}\,, \tag{9}$$

$$\mathrm{BIAS}(\Delta q_{\mathfrak{L}}) = \frac{\sum_{j=1}^{n_j} W_{\mathfrak{L},j}\,\Delta q_{\mathfrak{L},j}}{\sum_{j=1}^{n_j} W_{\mathfrak{L},j}}\,, \tag{10}$$

$$\mathrm{STD}(\Delta q_{\mathfrak{L}}) = \sqrt{[\mathrm{RMS}(\Delta q_{\mathfrak{L}})]^2 - [\mathrm{BIAS}(\Delta q_{\mathfrak{L}})]^2}\,. \tag{11}$$

where the water vapor weighting factor, $W_{\mathfrak{L},j}$, can assume one of three forms (W^0, W^1, or W^2)

$$W_{\mathfrak{L},j} = \begin{cases} 1 & , \mathrm{W}^0 \\ q_{\mathfrak{L},j} & , \mathrm{W}^1 \\ (q_{\mathfrak{L},j})^2 & , \mathrm{W}^2\,. \end{cases} \tag{12}$$

W^0 ($W_{\mathfrak{L},j} = 1$) simply reduces equations (9) and (10) to arithmetic means analogous to (3) and (4). To minimize skewing caused by dry atmospheres, sounder science team convention has been to use W^2 for RMS equation (9) and W^1 weighting in the attendant BIAS calculation equation (10). The reason for these weightings is to render both equations (9) and (10) in terms of the absolute differences, $\hat{q}_{L,j} - q_{L,j}$, instead of fractional or percent differences, $\Delta q_{\mathfrak{L},j}$ defined in (8). However, this combination will render RMS and BIAS metrics that are not, strictly speaking, compatible with one another. In most cases the incompatibility may not be significant, but under certain circumstances (e.g., polar analyses when there is a nonnegligible systematic error present), BIAS $\geq$ RMS can result. Because the Level 1 requirements (Table 2) for RMS were derived based upon statistics using the W^2 weighting, it is desirable that validation requirement assessments also use this weighting. However, regardless of what weighting is used for RMS, it is our recommendation that BIAS calculations use consistent weighting in future validation efforts. Given a correlative ozone sounding, the statistics are calculated for ozone in a manner similar to water vapor, namely equations (9) and (10), with W^2 weighting being desirable for both RMS and BIAS.

3.2. Use of Averaging Kernels

[20] *Rodgers* [1990] and *Rodgers and Connor* [2003] have shown the importance of using of averaging kernels (AKs) [*Backus and Gilbert*, 1970; *Conrath*, 1972] for performing proper intercomparisons of profiles obtained by two different observing systems and for characterizing retrieval uncertainty. The AK is defined as the Fréchet derivative of the retrieved state at layer L with respect to the "true" state at layer L' and is thus a square matrix with elements given by

$$A_{L,L'} \equiv \frac{\partial \widehat{X}_L}{\partial X_{L'}}, \tag{13}$$

where $\widehat{X}_L$ and $X_{L'}$ denote the retrieved and true states (i.e., T, H_2O, O_3, and other gas abundances) at layers L and L', respectively. In matrix form it will be denoted $\mathbf{A}$ and is dimensioned $m \times m$, where m is the number of retrieved layers. $\mathbf{A}$ may be thought of as a vertical "smoothing function" [*Rodgers*, 1990; *Rodgers and Connor*, 2003] that results from the finite resolution of passive sounders. The observing system AKs may be used for characterization of retrieval null-space errors defined by [*Rodgers*, 1990; *Rodgers and Connor*, 2003]

$$\epsilon_{\mathbf{s}} = (\mathbf{A} - \mathbf{I})(\mathbf{x} - \mathbf{x_0}), \tag{14}$$

where $\mathbf{I}$ is the identity matrix, $\mathbf{x}$ is the correlative state vector (e.g., RAOB reduced to forward model layers as described in Appendix B), and $\mathbf{x_0}$ is the a priori state vector or linearization point. The retrieval error analysis, $\hat{\mathbf{x}} - \mathbf{x}$, may be performed subtracting the null-space error as

$$(\hat{\mathbf{x}} - \mathbf{x}) - \epsilon_{\mathbf{s}} \;=\; \hat{\mathbf{x}} - (\mathbf{x} + \epsilon_{\mathbf{s}}) \;=\; \hat{\mathbf{x}} - \mathbf{x_s}, \tag{15}$$

where $\mathbf{x_s} \equiv \mathbf{x} + \boldsymbol{\epsilon}_{\mathbf{s}}$ is the correlative profile smoothed; from equation (14), $\mathbf{x_s}$ is calculated as

$$\mathbf{x_s} = \mathbf{A}(\mathbf{x} - \mathbf{x_0}) + \mathbf{x_0}. \tag{16}$$

The smoothed profiles $\mathbf{x_s}$ can then be used in equations (1), (2), (6), and (8) for descriptive statistics calculations as detailed in section 3.1. Because AKs factor into error characterization in this manner, basic recipes for their use in validation of EDR algorithms including legacy iterative methods (e.g., NUCAPS, NOAA Unique IASI, and AIRS), as well other physical methods (e.g., the original CrIMSS IDPS algorithm), are summarized below.

[21] Operational AK products developed by *Maddy and Barnet* [2008] are currently available for use as part of the standard AIRS, NOAA Unique IASI (NOAA-IASI), and NUCAPS (see section A1) operational product packages. These AK products can be utilized for smoothing correlative data using the method described by *Maddy and Barnet* [2008] which we summarize here. Within the AIRS algorithm (and subsequently the NUCAPS and NOAA-IASI algorithms), the retrieval space is reduced from the subset of n forward model layers above the bottom level l_b (defined in Appendix B), to m retrieved "coarser" layers, $\mathfrak{L}$, where $m < n \leq n_L = 100$ (note that the coarse retrieval layers $\mathfrak{L}$ are different from the coarse layers discussed in section 3.1). *Maddy and Barnet* [2008] point out that because trapezoidal basis functions, $\mathbf{F}$ (dimensioned $n \times m$), are used to map the m retrieval perturbations to forward model perturbations on the finer n forward model layers, smoothing thus occurs not only from $\mathbf{A}$ but also from $\mathbf{F}$. Given $\mathbf{F}$, an "effective averaging kernel," $\mathbf{A_e}$ ($n \times n$), is thus derived as [*Maddy and Barnet*, 2008]

$$\mathbf{A_e} = \mathbf{F}\,\mathbf{A}\,\mathbf{F}^{+}, \tag{17}$$

where $\mathbf{F}^{+}$ ($m \times n$) is the pseudoinverse of the trapezoidal basis functions given by

$$\mathbf{F}^{+} = (\mathbf{F}^T\,\mathbf{F})^{-1}\mathbf{F}^T, \tag{18}$$

superscript T denoting the transpose. Smoothing of the correlative profiles, $\mathbf{x}$ (reduced to forward model layers located above l_b, as described in Appendix B), is then achieved for temperature by substituting $\mathbf{A_e}$ for $\mathbf{A}$ in equation (16). For gas abundances (e.g., H_2O, O_3, and such), the smoothing is performed in logarithmic space by also substituting $\ln(\mathbf{x})$ for $\mathbf{x}$ and $\ln(\mathbf{x_0})$ for $\mathbf{x_0}$ in (16) [*Maddy and Barnet*, 2008].

[22] In contrast to legacy sounder methods using trapezoids, the original algorithm developed for the CrIMSS IDPS (see Appendix A2) is an OE retrieval performed on principal components (herein referred to as empirical orthogonal functions, EOFs) of the state vectors. The AKs are calculated in compressed EOF space ($m = 71$) as

$$\mathbf{A_c} = (\mathbf{K}^T\,\mathbf{S_y}^{-1}\,\mathbf{K} + \mathbf{S_{x_0}}^{-1})^{-1}\mathbf{K}^T\,\mathbf{S_y}^{-1}\,\mathbf{K}, \tag{19}$$

where $\mathbf{K}$ ($n_\nu \times m$, where $n_\nu = 1191$ channels) is the Jacobian matrix (i.e., the derivatives of the CrIS radiances with respect to state vectors, transformed to EOF space), $\mathbf{S_{x_0}}$ ($m \times m$) is the error covariance matrix associated with the a priori state vector, and $\mathbf{S_y}$ ($n_\nu \times n_\nu$) is the measurement error covariance matrix associated with CrIS/ATMS instrument noise and forward model errors. Because the algorithm calculates $\mathbf{A_c}$ ($m \times m$) in EOF space, the linear transformation to state space $\mathbf{A}$ ($n \times n$) is given by

$$\mathbf{A} = \mathbf{U}\,\mathbf{A_c}\,\mathbf{U}^{+}, \tag{20}$$

where $\mathbf{U}$ ($n \times m$) is the semiorthogonal matrix of EOFs and $\mathbf{U}^{+}$ ($m \times n$) is the pseudoinverse

$$\mathbf{U}^{+} = (\mathbf{U}^T\,\mathbf{U})^{-1}\mathbf{U}^T. \tag{21}$$

Because $\mathbf{U}^T\,\mathbf{U} = \mathbf{I}_m \Longrightarrow \mathbf{U}^{+} = \mathbf{U}^T$, equation (20) is reduced to

$$\mathbf{A} = \mathbf{U}\,\mathbf{A_c}\,\mathbf{U}^T, \tag{22}$$

which can be used to transform $\mathbf{A_c}$ derived from (19) back to state space. Smoothing of the layer-reduced correlative profiles, $\mathbf{x}$ for temperature and $\ln(\mathbf{x})$ for gas abundance, can then proceed as above via equation (16).

4. Application to CrIMSS

[23] This section provides an overview of the JPSS cal/val program and CrIMSS EDR validation efforts as they relate to the validation methodology discussed in section 2.

4.1. JPSS Cal/Val Program Overview

[24] The JPSS cal/val program is designed to ensure that the data products comply with the JPSS Level 1 accuracy requirements for both the AVTP and AVMP products as given in Tables 1 and 2, respectively. JPSS cal/val has been organized into four distinct phases that roughly correspond with the various maturity levels of the satellite

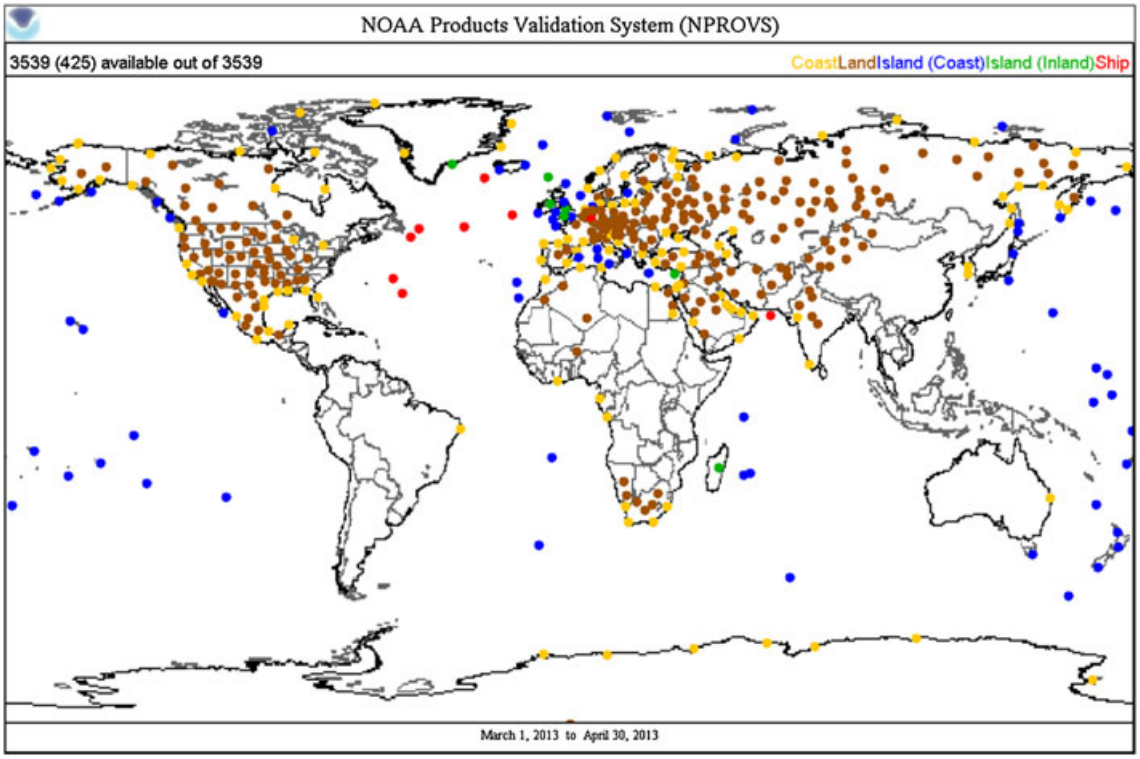

Figure 1. Locations of NPROVS conventional RAOBs collocated with S-NPP overpasses (3 h and 100 km window, March–April 2013). Matchup locations are color coded for surface types: coastal (yellow), land (brown), island (blue), inland island (green), and ship (red).

product lifetimes (including S-NPP, JPSS-1 and JPSS-2), namely the prelaunch, early orbit checkout (EOC), intensive cal/val (ICV), and long-term monitoring (LTM) phases described briefly below. Prelaunch phase efforts primarily support launch readiness. RTM simulations are first used to test the algorithm for theoretical robustness based upon controlled state parameters. Later, proxy data sets (i.e., data derived from existing satellite systems with similar specs) are utilized [e.g., *Xie et al.*, 2013]. The EOC phase encompasses roughly the first 90 days postlaunch as sensors are activated. EDR processing and validation activities commence when SDRs are available, at which point Beta-Maturity EDR (early release products) are compared against global numerical model analyses/forecasts (the first step in the validation hierarchy; see section 2.1) during initial Focus Days. The ICV phase begins after SDRs are stable and can continue up to 3+ years after launch for Provisional and Validated Maturity assessments. ICV may include additional model comparisons for subsequent Focus Days, as well as satellite EDR intercomparisons contingent upon the acquisition of simultaneous overpasses. RAOB matchup data sets (conventional, dedicated, and reference) are acquired during this period for conducting statistical assessments of the product. Contingent upon funding, intensive validation field campaigns ideally are conducted. The LTM phase extends from the end of the ICV phase until the end of the satellite operational lifetime, during which characterization of EDR products and long-term demonstration of performance is performed with respect to the hierarchal data sets described in section 2, especially conventional and dedicated RAOB assessments.

4.2. S-NPP Validation Data Sets

[25] This section summarizes the specific validation data sets for S-NPP CrIMSS ICV and LTM, tying them back to the validation hierarchy described in section 2. These include Focus Day data sets versus global NWP models, comparisons against conventional RAOBs and legacy products, and special comparisons against dedicated and reference RAOBs. As of this writing, an aircraft field campaign for S-NPP involving S-HIS and NAST-I was successfully completed in May 2013 out of Palmdale, California, USA, with the data analysis currently ongoing. Detailed results from all the above platforms will be subjects of future work.

4.2.1. Global Focus Days

[26] Global data samples have been acquired from Focus Days which include model (viz., ECMWF) and EDR matchups from other satellites and/or algorithms. Both the IDPS EDR and the NUCAPS EDR have been directly compared against one another to demonstrate algorithm differences, and these have been compared against different sensors on different platforms, including AIRS and COSMIC. Comparisons against AIRS and COSMIC are given in the related EDR paper in this Special Section [*Divakarla et al.*, 2014]. AIRS observations are uniquely suited for intercomparisons with S-NPP due to the Aqua local equator crossing time coinciding with that of S-NPP, allowing for simultaneous nadir overpasses. COSMIC data are well suited to providing reliable information in the upper troposphere and stratosphere where radiosondes and numerical models have known limitations [e.g., *Sun et al.*, 2010, 2013; *Seidel et al.*, 2011]. The methodology

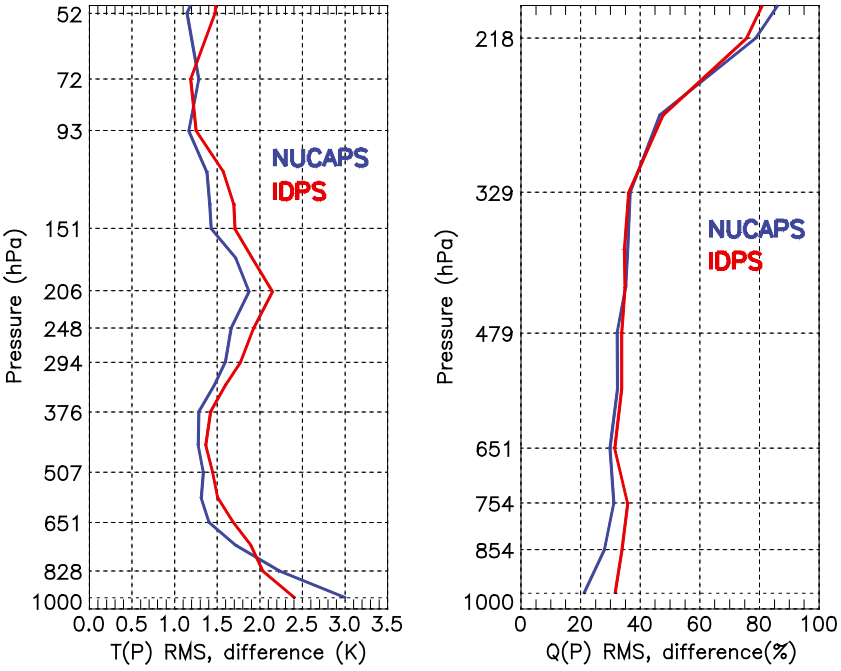

Figure 2. NPROVS-derived preliminary RMS statistics of CrIMSS NUCAPS (blue) and original IDPS (red) EDR relative to conventional RAOB matchups: (left) AVTP and (right) AVMP. The matchup sample used (cf. Figure 1) includes only accepted IR + MW retrievals by both the NUCAPS and IDPS systems (n = 3750).

for the comparison of COSMIC temperature profiles and sounder retrievals is the subject of a future paper (M. Feltz et al., Methodology for the validation of temperature profiles from hyperspectral infrared sounders using GPS radio occultation: Experience with AIRS and COSMIC, *Journal of Geophysical Research: Atmospheres*, submitted manuscript, 2013).

4.2.2. Conventional RAOBs and Legacy Soundings

[27] The NOAA Products Validation System (NPROVS) is a near-real-time satellite validation system developed at NESDIS/STAR that provides a monitoring and validation capability for intercomparing sounder AVTP/AVMP EDR products (from multiple platforms and algorithms) against collocated conventional RAOBs and NWP models [*Reale et al.*, 2012b]. Details about the RAOB satellite collocation method and characteristics of radiosonde data including its atmospheric profile structure and measurement quality and accuracy can be found in *Reale et al.* [2012b] and *Sun et al.* [2010, 2013]. Ongoing assessments of CrIMSS sounding performance are being carried out under a variety of sampling conditions, including time and distance window, terrain, EDR and RAOB quality flags, cloud condition, and radiosonde types.

[28] Figure 1 shows the locations of NPROVS conventional RAOB collocations (3 h, 100 km matchup window) with S-NPP acquired over the months of March through April 2013. Although there is representation among different zones, the sample is skewed toward continental, midlatitude coastal and land locations. Figure 2 shows the corresponding RMS difference of the NUCAPS and the original IDPS retrievals (provisional maturity status) from the RAOB matchup sample. It can be seen that both EDR algorithms are performing comparably.

4.2.3. Dedicated and Reference RAOBs

[29] A number of JPSS supported ICV phase dedicated RAOB sites have been coordinated, leveraging a number of collaborating institutions. Figure 3 and Table 3 summarize the Year 1 dedicated RAOB data, and a brief description of each site is given below. Similar data sets are planned for acquisition during Years 2 and 3. Unless otherwise specified, all dedicated RAOBs have been acquired using Vaisala RS92-SGP radiosondes launched approximately 15–30 min prior to S-NPP overpass times and processed with the latest Vaisala DigiCORA-III software (v3.64).

[30] Reference RAOBs from GRUAN sites have been routinely utilized at STAR in support of product validation since 2011. Most GRUAN RAOBs are launched at synoptic times although at some sites, there are launches that meet the synchronization requirements of dedicated RAOB coincidence with respect to satellite overpass. These are considered a key component of the validation effort particularly as they broaden the global distribution and provide traceable reference profiles that include profile uncertainty estimates [*Immler et al.*, 2010]. Reprocessing dedicated RAOBs using GRUAN processing software is desirable as it would elevate these to traceable reference data.

4.2.3.1. PMRF Site

[31] The first S-NPP dedicated RAOB data set (which provided a "test bed" for the early validation effort) was collected by The Aerospace Corporation (hereafter "Aerospace") for overpasses of the Aerospace Transportable Lidar System-2 [e.g., *Wessel et al.*, 2008] resident at the

Table 3. Dedicated RAOB Sites (Year 1)[a]

	ARM-TWP[c]	*ARM-SGP*	*ARM-NSA*	*PMRF*	*BCCSO*[b]	*AEROSE*
Location	Manus Island, Papua New Guinea	Ponca City, Oklahoma, USA	Barrow, Alaska, USA	Kauai, Hawaii, USA	Beltsville, Maryland, USA	Tropical Atlantic Ocean
Regime	Tropical Pacific Warm Pool, Island	Midlatitude Continent, Rural	Polar Continent	Tropical Pacific, Island	Midlatitude Continent, Urban	Tropical Atlantic SAL/Dust, Smoke
The n planned	90	90 × 2	90 × 2	40	ad hoc	≃60–100
The N acquired	94	95 + 93	95 + 91	40	31	69
The N_{BE} acquired	80	89	90	—	—	—
Sampling period	Aug 2012 to Jun 2013	Jul–Dec 2012	Jul–Dec 2012	May, Sep 2012	Jun 2012 to Feb 2013	Jan–Feb 2013

[a]Year 1 (2012–2013) S-NPP dedicated RAOBs, planned number (n), and acquired number (N). The ARM-SGP (Southern Great Plains) and North Slope of Alaska (NSA) sites launched two RAOBs per overpass (45 and 5 min prior to overpass) for site state best estimates (N_{BE}). The planned September 2012 Aerosols and Ocean Science Expeditions (AEROSE) campaign was postponed to January–February 2013, but two were launched in September 2012 and one in March 2013, during the transits to and from Charleston, SC. Similar numbers of dedicated radiosondes for ARM, AEROSE, and Pacific Missile Range Facility (PMRF) have been secured for Year 2 as of this writing; dedicated radiosondes are also planned for Year 3 and beyond, but these are contingent upon funding.
[b]BCCSO, Beltsville Center for Climate System Observation.
[c]TWP, Tropical Western Pacific.

Pacific Missile Range Facility (PMRF), Barking Sands, Hawaii (22.052°N, 159.782°W; elevation 2 m) [*Mollner et al.*, 2013]. In addition to RS92 radiosondes, a Raman lidar measured profiles of H_2O (1–20 km) and temperature (30–60 km) during nighttime overpasses [*Mollner et al.*, 2013], and merged profiles were generated by combining both systems, including estimated uncertainties in the data. In 2012 (Year 1), data were collected for a total of 40 NPP overpasses (elevation angles ≥ 40° and without dense cloud cover) during May and September. In 2013 (Year 2) another data collection campaign was completed in May–June for a total of 36 overpasses.

[32] Figure 4 shows the initial test bed results comparing NUCAPS EDR product retrievals against PMRF May 2012 RAOBs for field-of-regard (FOR) within 150 km of the launch location. RMS and BIAS $\pm 1\sigma$ statistics are calculated according to the methodology described in section 3.1 and Appendix B, with water vapor W^2 weighting defined by equation (12). For this small tropical maritime sample ($n = 105$), the NUCAPS AVTP RMS (top left) is seen to meet the tropospheric requirements defined in Table 1, with the AVMP RMS close to meeting lower tropospheric requirements defined in Table 2. Similar results for the original IDPS Provisional Maturity EDR are presented in Figure 15 of *Divakarla et al*. [2014].

4.2.3.2. Beltsville Center for Climate System Observation Site

[33] The Howard University Beltsville Center for Climate System Observation (BCCSO) research site is located at 39.0542°N, 76.8775°W at an elevation of 52.8 m above sea level in Beltsville, Maryland (nearby NESDIS/STAR in College Park). The site is part of the initial GRUAN candidate sites (see Figure 5) and began supporting launches of Vaisala RS92 dedicated radiosondes coincident with selected S-NPP satellite overpasses in 2012. The launches have preferentially targeted high-elevation clear-sky

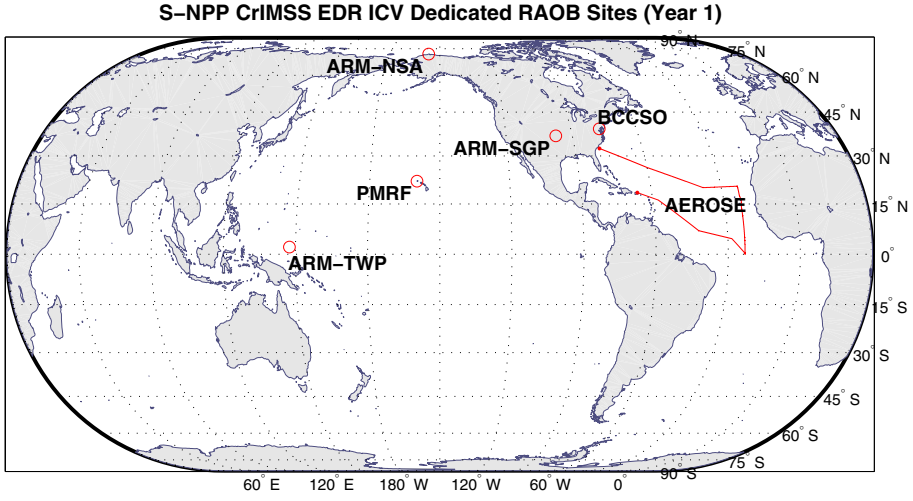

Figure 3. Locations of the participating S-NPP CrIS/ATMS EDR Intensive cal/val (ICV) dedicated RAOB sites for Year 1 (2012–2013). The red line shows the cruise track of the NOAA Ship *Ronald H. Brown* during AEROSE, 8 January to 13 February 2013. Year 2 sites are identical with the exception of the AEROSE cruise track. Map projection is equal area.

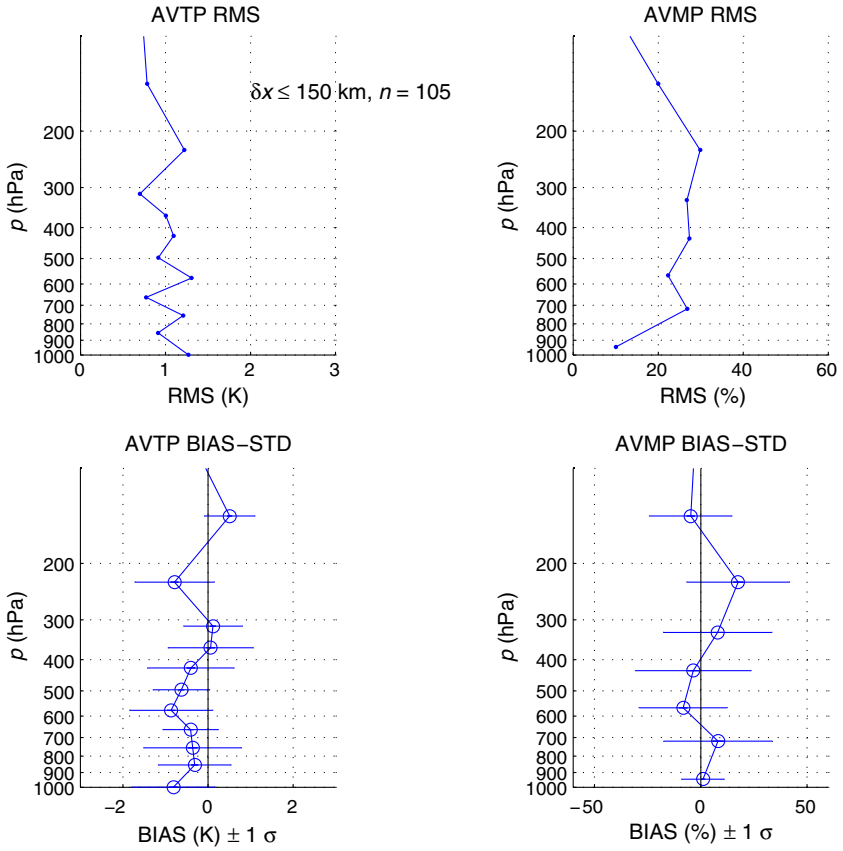

Figure 4. Preliminary assessment of NUCAPS EDR products versus dedicated RAOB matchups (for FOR within $\delta x \leq 150$ km of the PMRF RAOB launch location, $n = 105$) launched from the PMRF test bed site during May 2012. The left and right columns are for AVTP and AVMP, respectively, with the top and bottom rows showing RMS and BIAS $\pm 1\sigma$, respectively. RMS, BIAS, and STD (σ) statistics are calculated using the methodologies detailed in section 3.1 and Appendix B.

overpasses, and these are supplemented with Raman lidar and MWR remotely sensed time-height T and H_2O profiles. Future plans include inclusion of a CFH launch at a frequency of one launch per month. A primary objective will be to provide reference-quality RAOBs by applying the GRUAN processing and correction algorithm (see section 4.2.3.4 below). Together, the GRUAN-corrected RS92 profiles and the Raman/MWR measured time-height profiles will provide best estimates of the atmospheric state at the time of overpass.

4.2.3.3. ARM Sites

[34] The Year 1 ARM validation effort involved launching Vaisala RS92 radiosondes from the Southern Great Plains (SGP), North Slope of Alaska (NSA), and Tropical Western Pacific (TWP) Manus island sites (for more on ARM, see *Mather and Voyles* [2013] and *Long et al.* [2013]). After the initial checkout of the ATMS and CrIS SDRs, the dedicated ARM launches commenced in July 2012. This effort included two launches per overpass at the SGP and NSA sites, and one launch per overpass at the TWP Manus site, with the goal of collecting data for a total of 90 overpass matchups at each site (see Table 3). A second phase of radiosonde launches (Year 2) at the same three sites commenced in June 2013.

[35] These ARM sites represent three very different climatic regimes with a range of conditions and retrieval difficulty. Further science justification and details of the approach for this effort are described in detail in *Tobin et al.* [2006a]. Each RAOB is scaled such that the total column precipitable water vapor is equal to that measured by a ground-based microwave radiometer (MWR). These scaled RAOBs are then interpolated to the overpass time and combined with other ancillary ARM data to produce a "best estimate" of the atmospheric state for the overpass times at each site.

4.2.3.4. GRUAN Sites

[36] GRUAN sites currently include 19 active plus three future sites shown in Figure 5. Not all the sites are actively producing routine RS92 RAOB observations, but all sites are expected to begin providing observations in the future. Since 2010, over 20,000 reference RAOBs are available from the combined sites with over half of them deemed suitable reference data by the GRUAN Lead Center (LC) at Deutscher Wetterdienst, Lindenberg, Germany. All of these observations, including those which did not initially pass the GRUAN reference criteria, are stored at STAR/NPROVS. In addition to using the data in S-NPP validation, STAR also provides feedback to the LC, for example, on the respective performance of passed versus rejected GRUAN data in the context of validation and whether some restrictions may be too strict (or not strict enough). A key performance parameter from GRUAN is the uncertainty estimate which provides

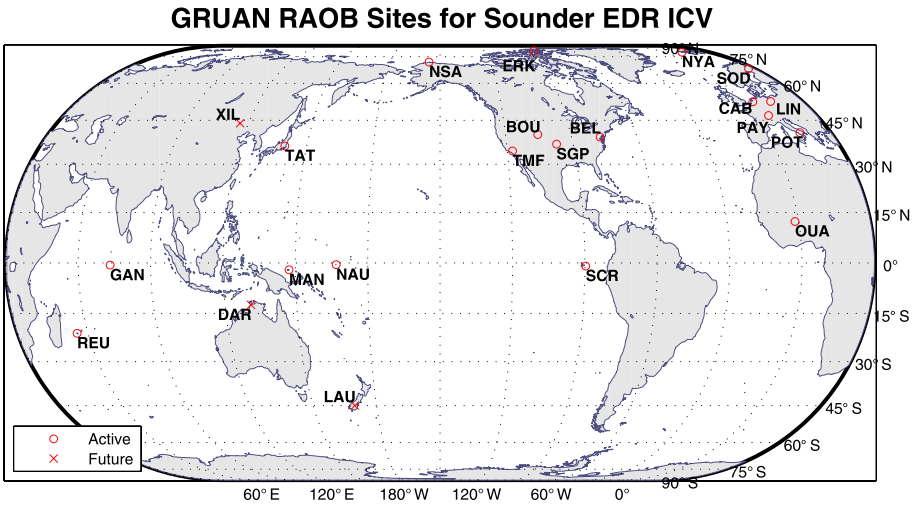

Figure 5. Locations of GRUAN sites (active and future), providing reference RAOBs for satellite sounder EDR ICV via matchups acquired by the NPROVS$^+$ system. Three-letter site identifiers are as follows (pending sites denoted with asterisk): LIN (Lindenberg, Germany); BEL (Beltsville, Maryland, USA); BOU (Boulder, Colorado, USA); CAB (Cabauw, The Netherlands); DAR (Darwin, Australia) ERK (Eureka, Canada); GAN (Gan, Maldives Islands); LAU (Lauder, New Zealand); MAN (Manus Island, Papua New Guinea); NAU (Republic of Nauru); NSA (Barrow, Alaska, USA); NYA (Ny-Ålesund, Norway); OUA (Ouagadougou, Burkina Faso); PAY (Payerne, Switzerland); POT (Potenza, Italy); REU (La Reunion Island); SCR (San Cristobal, Galapagos Islands); SGP (Lamont, Oklahoma, USA); SOD (Sodankylä, Finland); TAT (Tateno, Japan); TMF (Table Mountain Facility, California, USA); XIL (Xilin Hot, China). Map projection is equal area.

an alternative approach for quantifying (and qualifying) S-NPP product performance. Plans to increase the GRUAN to up to 40 sites by 2016 are in the planning stages.

4.2.4. AEROSE Campaigns

[37] The NOAA Aerosols and Ocean Science Expeditions (AEROSE) are a series of trans-Atlantic field campaigns that have been conducted onboard the NOAA Ship *Ronald H. Brown* [*Morris et al.*, 2006; *Nalli et al.*, 2011] on a near-yearly basis, with eight campaigns (as of this writing) comprised of nine 4–6 week cruise legs taking place in 2004, 2006–2011, and 2013. The science focus of AEROSE is the observation of dust and smoke aerosols over the Atlantic, and it is out of this primary objective that satellite sounder validation has become a critical mission component [*Nalli et al.*, 2006, 2011]. Within the validation hierarchy (section 2), AEROSE (and to a limited extent, the PMRF and BCCSO sites) have become important campaigns of opportunity (section 2.5) for S-NPP validation.

[38] AEROSE campaigns offer several distinct advantages for satellite validation [*Nalli et al.*, 2006, 2011]. First, ocean-based measurements are advantageous because the sea surface is radiatively well-characterized in terms of emissivity and skin temperature [e.g., *Smith et al.*, 1996; *Minnett et al.*, 2001; *Nalli et al.*, 2008], and oceanic regions are where satellite data have the greatest positive impact on NWP. Furthermore, dedicated RAOB data acquired during AEROSE are not uploaded into the Global Telecommunications System (GTS). Because these data are not assimilated and because the AEROSE domain is far removed from any land-based GTS RAOBs, AEROSE RAOBs constitute a genuinely independent truth data set [*Nalli et al.*, 2011]. Additionally, AEROSE campaigns also acquire dedicated ozonesonde and concurrent M-AERI skin temperature measurements. These data sets combine to allow a near complete specification of the atmospheric state and thus open the possibility of "dissections" as discussed in section 2.5.

[39] The first AEROSE campaign supporting ICV phase S-NPP CrIMSS validation was successfully conducted in January–February 2013 (5.5 weeks) (see Figure 3 for cruise track). A total of 67 (23) sondes (ozonesondes) were launched for S-NPP, along with 43 (1) for MetOp. Several RAOB matchups were obtained during dust and smoke aerosol outflows (aerosol optical depths reaching 0.6) under clear to partly cloudy sky conditions, including an ideal observing period $\simeq$26 January 2013 where there were no visible clouds within any lines of sight. Land-based RAOB launches were coordinated with the Puerto Rico National Weather Service and the University of Puerto Rico Mayagüez during a coincident S-NPP overpass for comparisons of RAOB types: island based (conventional and dedicated) and ship based. Analyses of these data, including ozone profile validation, are subjects of future work. The Year 2 AEROSE campaign was successfully completed in November–December 2013 (Barbados to Brazil).

4.3. Future Efforts

[40] The EDR validation methodology, statistical metrics, and data sets discussed in this paper are intended as a standard reference for current (e.g., S-NPP, MetOp) and future (e.g., JPSS-1 and JPSS-2) satellite sounder validation efforts. S-NPP CrIMSS efforts are ongoing and we have highlighted some initial results above. Details on the Beta and Provisional Maturity assessments of the original operational IDPS CrIMSS product can be found in the related EDR paper [*Divakarla et al.*, 2014] in this Special Section. Detailed validation assessments

using EDR intercomparisons and RAOBs (both conventional and dedicated/reference) for the Stage 1–3 Validated Maturity assessments and LTM phase will be the subjects of future work.

Appendix A: CrIMSS Operational EDR Products

[41] There have been two CrIMSS retrieval algorithms implemented operationally, both being designed to address the requirements on accuracy, latency, and yield. These are briefly summarized below.

A1. NUCAPS EDR Algorithm

[42] The NUCAPS algorithm (A. Gambacorta et al., manuscript in preparation, 2013) is a multistep iterative method (regularized least squares minimization) that solves for parameters (AVTP, AVMP, trace gas abundances, surface temperature, cloud parameters, emissivity, and such) in separate steps while holding other variables constant. This system processes the CrIMSS data in a manner similar to that obtained from heritage AIRS/AMSU [*Aumann et al.*, 2003; *Chahine et al.*, 2006] and NOAA-unique IASI/AMSU [*Cayla*, 1993] systems, with the EDR retrieval algorithm being an exact line-for-line modular implementation of the iterative, multistep AIRS Science Team retrieval algorithm [*Susskind et al.*, 2003, 2011]. The NUCAPS retrieval algorithm consists of the following steps. First, a MW retrieval produces cloud liquid water flags and microwave surface emissivity uncertainty. This is followed by a fast eigenvector regression retrieval for temperature and moisture [e.g., *Goldberg et al.*, 2003] that is trained using the ECMWF analysis and the warmest FOV in the CrIS/ATMS field-of-regard (FOR) (an array of 3×3 CrIS FOVs collocated with an ATMS FOV). The cloud-clearing algorithm then uses a selected set of IR channels [*Gambacorta and Barnet*, 2013] to retrieve the IR CCR (clear column) product and reject cases violating the cloud-clearing requirements. A second fast eigenvector regression retrieval is then performed [e.g., *Goldberg et al.*, 2003] on the ATMS and CCRs to produce the final regression-based temperature and moisture profiles. This second regression retrieval is finally used as an initial solution for the physical retrieval algorithm based on an iterated regularized least squared minimization applied to a selected set of IR channels [*Gambacorta and Barnet*, 2013], starting with the AVTP retrieval (temperature being the most linear component of the radiative transfer equation), followed by AVMP and operational trace gas profiles: O_3, CO, CO_2, CH_4, N_2O, HNO_3, and SO_2.

A2. Original IDPS EDR Algorithm

[43] The original CrIMSS EDR algorithm (developed by AER and NGAS) [*Moncet et al.*, 2005] uses a classical constrained inversion OE method to estimate simultaneously the geophysical state of both the atmosphere (AVTP, AVMP, AVPP EDR) and the surface [*Lynch et al.*, 2009]. The first guess is derived from an ATMS MW-only physical retrieval. In the first stage, MW-only retrievals are performed on the CrIS/ATMS FOR to obtain an estimate of the CCR spectrum in a manner similar to *Susskind et al.* [2003]. These are then used in the second stage for the simultaneous retrieval of atmospheric and surface states under clear to partly cloudy (MW + IR) or cloudy (MW only) conditions. The simultaneous retrieval is attained by performing an inversion in principal component space using an OE approach that minimizes a cost function on the CrIS/ATMS FOR. Cloud clearing is a key component of this second-stage processing, and accuracy of the cloud-cleared IR radiance spectrum determines the final quality of the output. More on the algorithm theoretical basis can be found in *Moncet et al.* [2005] and *Lynch et al.* [2009]; more on the offline implementation and testing of the original operational IDPS EDR product can be found in *Divakarla et al*. [2014].

Appendix B: Reduction to Correlative Layers

[44] As mentioned in section 3.1, proper comparisons between a directly measured high-resolution state (i.e., "truth" obtained independent of the satellite observing system, e.g., in situ RAOB) and a satellite-retrieved state require that the measured state be reduced to effective layer quantities to yield a correlative true state. One approach for doing this might simply be to take binned means for layers defined by boundaries throughout the column. However, we propose and utilize a molecular-conservative approach that is physically analogous to the forward model used in the retrieval [e.g., *Rodgers*, 1990]

$$\hat{\mathbf{x}} = I[F(\mathbf{x}, \mathbf{b}), \mathbf{b}, \mathbf{c}], \tag{B1}$$

where $\mathbf{x}$ and $\hat{\mathbf{x}}$ are the true and retrieved state vectors dimensioned $1 \times n_L$ (T, q, O_3, and such), respectively, F is the forward model with parameters $\mathbf{b}$ (e.g., spectroscopy), and I is the inverse model (i.e., retrieval), with parameters $\mathbf{c}$ not included in F (i.e., unrelated to the measurement) [*Rodgers*, 1990]. Radiative transfer within F [e.g., *Susskind et al.*, 1983; *Strow et al.*, 2003] involves a path integral through an attenuating medium which is approximated via finite layer "slabs" [e.g., *Armstrong*, 1968; *Gallery et al.*, 1983; *Stephens*, 1994]. To obtain the correlative layer values from high-resolution measured quantities, we similarly integrate molecular column density over the atmospheric column, interpolate those to forward model layer boundaries (i.e., levels), then compute effective layer quantities from the interpolated level values. This approach essentially accumulates the atmospheric path as a function of altitude thereby allowing it to be redivided into arbitrary "slices" defined by F. The details of this approach along with their theoretical basis are formulated below.

[45] Effective layer pressures, $\bar{p}_L$, are defined as follows:

$$\bar{p}_L \equiv \frac{\int_{z_{l+1}}^{z_l} p(z)\,\mathrm{d}z}{\int_{z_{l+1}}^{z_l} \mathrm{d}z} = \frac{1}{z_l - z_{l+1}} \int_{z_{l+1}}^{z_l} p(z)\,\mathrm{d}z, \tag{B2}$$

where z is geopotential height and subscripts l denote layer boundaries (levels). The hypsometric equation is given by [e.g., *Wallace and Hobbs*, 1977]

$$z_l - z = H \ln\left(\frac{p}{p_l}\right), \tag{B3}$$

where H is the scale height of the layer. Equation (B3) is also equivalently expressed as

$$p(z) = p_l \exp\left(\frac{z_l - z}{H}\right) . \quad \text{(B4)}$$

Integrating (B4) over a thin layer (where T is assumed constant) and substituting yield

$$\int_{z_{l+1}}^{z_l} p(z)\,\mathrm{d}z = H p_l \cdot \left[\exp\left(\frac{z_l - z_{l+1}}{H}\right) - 1\right] \quad \text{(B5)}$$

$$= H \cdot (p_{l+1} - p_l) . \quad \text{(B6)}$$

While any number of layers n_L may be used within an arbitrary forward model F, *Strow et al.* [2003] found $n_L = 100$ to be a sufficient number and thus defined a standard set of $n_l = n_L + 1 = 101$ levels for the CrIS/IASI/AIRS rapid transmittance algorithm (RTA). The n_L forward model layer pressures can be derived by substituting equations (B3) and (B6) into (B2), which results in

$$\overline{P}_L = \frac{P_{l+1} - P_l}{\ln(P_{l+1}/P_l)} , \quad l = L = 1, 2, \ldots, n_L , \quad \text{(B7)}$$

where uppercase P denote forward model pressures and subscripts L and l denote layer and level, respectively.

[46] For truth measurements or observations, we shall use lowercase p to denote pressure and subscripts $\mathcal{L}$ and ℓ to denote fine layer and level, respectively. Generally speaking, such data are at higher vertical resolution than the forward model, with the number of data points much greater than the forward model pressure grid (i.e., $n_\ell \gg n_l$). Given a typical radiosonde measuring pressure, temperature, and humidity, the number density (molecules/cm^3) for air is given (in centimeter-gram-second units) simply by [e.g., *Wallace and Hobbs*, 1977]

$$N_{a,\ell}(p_\ell, T_\ell) = 10^3 \frac{p_\ell}{k\,T_\ell} , \quad \ell = 1, 2, \ldots, n_\ell , \quad \text{(B8)}$$

where k is the Boltzmann constant in ergs, p_ℓ and T_ℓ are pressure (hPa) and temperature (K) measured at observation level ℓ, and the 10^3 factor converts pressure from hPa to dPa. Number densities for moisture and ozone, $N_{w,\ell}$ and $N_{o,\ell}$ (molecules/cm^3), are calculated from the radiosonde measurements of relative humidity (RH) % and O_3 partial pressure (mPa) (given an ozonesonde). For ozone the total atmospheric pressure in (B8) is replaced with the measurement of O_3 partial pressure (in mPa) and multiplied by a factor of 10^{-2} to convert to dPa. Similarly, number density for water vapor can be obtained by replacing the total pressure p_ℓ in equation (B8) by the water vapor partial pressure, e_ℓ. The vapor pressure can be calculated directly from RAOB RH as [e.g., *Stull*, 2000]

$$e_\ell = e_s(T_\ell) \frac{\mathrm{RH}}{100\%} , \quad \text{(B9)}$$

where e_s is the saturation vapor pressure (SVP) for temperature T_ℓ, or alternatively from H_2O mass mixing ratio, r_ℓ (a meteorological moisture parameter and conservative quantity), as

$$e_\ell = p_\ell \cdot \left(\frac{r_\ell}{r_\ell + \epsilon}\right) , \quad \text{(B10)}$$

where ϵ is the ratio of the molecular masses of water vapor and dry air, $\epsilon \equiv M_w/M_d \approx 0.622$. Because radiosondes typically measure RH, both equations (B9) and (B10) require the calculation of SVP (e_s), and we note here that there are a number formulas available for doing this. Nevertheless, *Miloshevich et al.* [2006] have recommended that the formulations of *Wexler* [1976] and *Hyland and Wexler* [1983] be used, both for their accuracy and for consistency with Vaisala radiosondes which also use those formulations. From equations (B8) and (B10), the H_2O number density is then given by

$$N_{w,\ell}(p_\ell, T_\ell, r_\ell) = 10^3 \frac{p_\ell}{(1 + \epsilon/r_\ell) k\,T_\ell} , \quad \ell = 1, 2, \ldots, n_\ell . \quad \text{(B11)}$$

[47] To account correctly for atmospheric path length, we first obtain column abundances for atmospheric species x, denoted Σ_x (molecules/cm^2), by integrating from the top measurement z_t (e.g., RAOB balloon burst altitude) to the measurement level height z as [e.g., *Armstrong*, 1968; *Gallery et al.*, 1983]

$$\Sigma_x(z) \equiv \int_{z_t}^{z} N_x(z')\,\mathrm{d}z' . \quad \text{(B12)}$$

Equation (B12) may be numerically computed via finite differencing of $n_\mathcal{L} = n_\ell - 1$ fine-layer mean molecular abundances, $\overline{N}_{x,\mathcal{L}}\,\delta z_\mathcal{L}$, with geopotential thicknesses $\delta z_\mathcal{L} \equiv z_{\ell+1} - z_\ell$ ($\ell = \mathcal{L} = 1, 2, \ldots, n_\mathcal{L}$) as

$$\Sigma_x(z) \approx \Sigma_{x,\mathcal{L}} \equiv \sum_{\mathcal{L}}^{n_\mathcal{L}} \overline{N}_{x,\mathcal{L}}\,\delta z_\mathcal{L} , \quad \text{(B13)}$$

where $\overline{N}_{x,\mathcal{L}} \approx (N_{x,\ell+1} + N_{x,\ell})/2$ and $(N_{x,\ell+1} \wedge N_{x,\ell}) > 0$.

[48] In a similar manner, the temperature profile is integrated weighted by the air number density [e.g., *Gallery et al.*, 1983] given by equation (B8) as

$$\Sigma_\mathcal{T}(z) \equiv \int_{z_t}^{z} T(z')\,N_a(z')\,\mathrm{d}z' , \quad \text{(B14)}$$

which is calculated as

$$\Sigma_\mathcal{T}(z) \approx \Sigma_{T,\mathcal{L}} \equiv \sum_{\mathcal{L}}^{n_\mathcal{L}} \overline{T}_\mathcal{L}\,\overline{N}_{a,\mathcal{L}}\,\delta z_\mathcal{L} , \quad \text{(B15)}$$

where $\overline{T}_\mathcal{L} \approx (T_{\ell+1} + T_\ell)/2$, $\overline{N}_{a,\mathcal{L}} \approx (N_{a,\ell+1} + N_{a,\ell})/2$ and $(N_{a,\ell+1} \wedge N_{a,\ell}) > 0$.

[49] Equations (B13) and (B15) can then be linearly interpolated to the truncated forward model pressure vector $\mathbf{P}$ (spanning the measured column) and simply denoted $\Sigma_{x,\mathcal{L}}(\mathbf{P})$ and $\Sigma_{T,\mathcal{L}}(\mathbf{P})$, respectively. $\mathbf{P}$ is defined as forward model (e.g., RTA) pressure levels P_l, including the observed surface and top levels. The top level l_0 is defined simply as the level just below the top observed pressure level, p_0 (e.g., at balloon burst altitude). Because of variable terrain and surface pressure, the bottom level, l_b, is less straightforward. A conventional approach has been to define l_b such that the full RTA layer $[P_{l_b-1}, P_{l_b}]$ is included in the calculation for a surface layer greater than 5 hPa in thickness, that is, l_b is taken to be the level where $p_s - P_{l_b-1} \geq +5$ hPa or equivalently $p_s - P_{l_b} < +5$ hPa (this follows since $p_s \geq P_{l_b} > P_{l_b-1} \Longrightarrow p_s - P_{l_b-1} > p_s - P_{l_b}$). The reason for this approach lies in the treatment of the forward calculation used in the retrieval algorithm and is employed here for consistency in the conversion to coarse layers for statistical calculations as described in section 3.1. In this way, the bottom layer usually encompasses p_s, that is $P_{l_b-1} < p_s < P_{l_b}$, except when

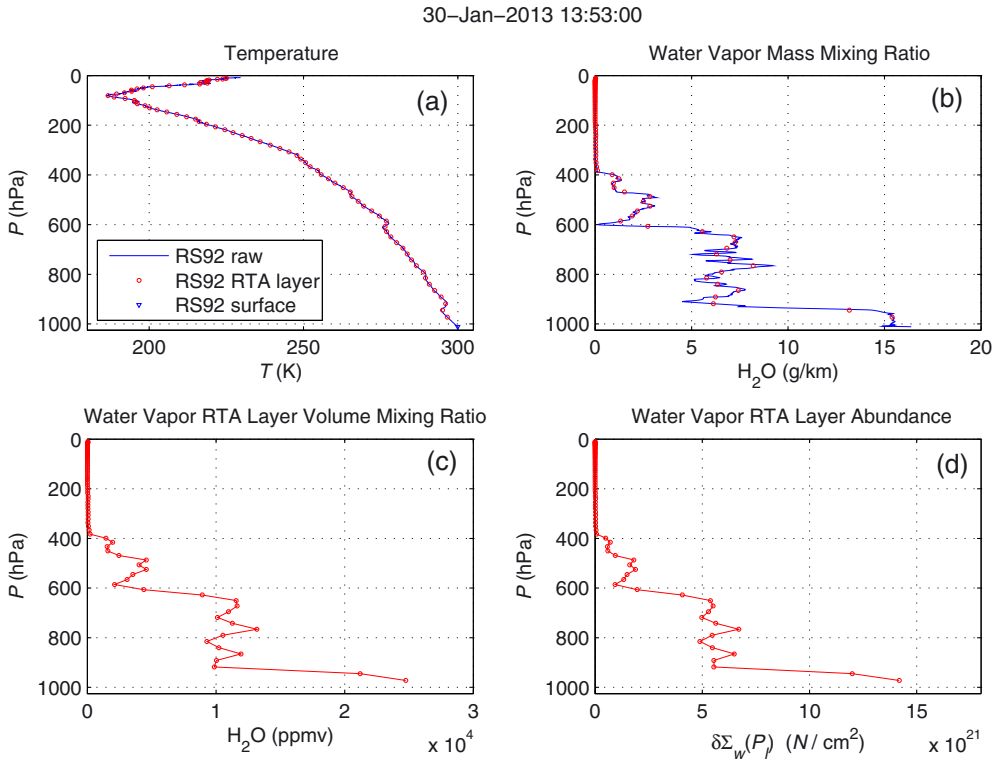

Figure B1. Example of high-resolution RAOB reduced to 100 effective RTA layers for an RS92 radiosonde launched during AEROSE on 30 January 2013 over the tropical Atlantic Ocean (5.0352°N, 23.0392°W): (a) RAOB and effective layer temperature using equation (B17), (b) RAOB and effective layer H_2O mass mixing ratio using (B18), (c) effective layer H_2O volume mixing ratio, and (d) effective layer H_2O abundance.

$0 < p_s - P_{l_b} < 5$ hPa, then $p_s > P_{l_b}$, and the interpolation vector **P** is thus given by

$$\mathbf{P} \equiv \begin{cases} [p_t, P_{l_t}, P_{l_t+1}, \ldots, P_{l_b-1}, P_{l_b}, p_s] \ , \ 0 < p_s - P_{l_b} < 5 \\ [p_t, P_{l_t}, P_{l_t+1}, \ldots, P_{l_b-1}, p_s, P_{l_b}] \ , \ p_s - P_{l_b} \geq 5 \, . \end{cases} \quad \text{(B16)}$$

[50] Given $\Sigma_{x,\mathcal{L}}(\mathbf{P})$ and $\Sigma_{T,\mathcal{L}}(\mathbf{P})$, we can now calculate the effective layer quantities. The effective layer pressures $\overline{P}_L$ for the sounding are simply computed as in (B7) except using the levels defined in (B16). Given $\Sigma_{a,\mathcal{L}}(\mathbf{P})$ and $\Sigma_{T,\mathcal{L}}(\mathbf{P})$, the effective layer temperatures $T_L(\overline{P}_L)$ are computed from

$$T_L(\overline{P}_L) = \frac{\Sigma_{T,\mathcal{L}}(P_{l+1}) - \Sigma_{T,\mathcal{L}}(P_l)}{\Sigma_{a,\mathcal{L}}(P_{l+1}) - \Sigma_{a,\mathcal{L}}(P_l)} \, , \quad l = L = 1, 2, \ldots, m_L \, , \quad \text{(B17)}$$

where m_L equals the length of vector **P**. Given $\Sigma_{x,\mathcal{L}}(\mathbf{P})$, and $\overline{T}_L(\overline{P}_L)$ from (B17), effective layer water vapor mass mixing ratio (g/kg relative to dry air) is calculated as

$$r_{w,L}(\overline{P}_L) = \epsilon \cdot 10^3 \frac{\Sigma_{w,\mathcal{L}}(P_{l+1}) - \Sigma_{w,\mathcal{L}}(P_l)}{\overline{N}_{a,L} \, \delta z_L - \left[\Sigma_{w,\mathcal{L}}(P_{l+1}) - \Sigma_{w,\mathcal{L}}(P_l)\right]} \, , \quad \text{(B18)}$$

where the atmospheric mean layer abundance is given by

$$\overline{N}_{a,L} \, \delta z_L \equiv 10^3 \frac{\overline{P}_L}{k \, \overline{T}_L(\overline{P}_L)} \, \delta z_L \, , \quad \text{(B19)}$$

and $\delta z_L \equiv z(P_l) - z(P_{l+1})$. Similarly, the effective layer ozone volumetric mixing ratio (ppbv, dry air) is given by

$$r_{o,L}(\overline{P}_L) = 10^9 \frac{\Sigma_{o,\mathcal{L}}(P_{l+1}) - \Sigma_{o,\mathcal{L}}(P_l)}{\overline{N}_{a,L} \, \delta z_L - \left[\Sigma_{w,\mathcal{L}}(P_{l+1}) - \Sigma_{w,\mathcal{L}}(P_l)\right]} \, . \quad \text{(B20)}$$

[51] Figure B1 shows an example RS92 RAOB that has been reduced to 100 RTA layers using the above methodology. The RAOB was acquired during the 2013 AEROSE campaign (see section 4.2.4) over the tropical Atlantic Ocean (5.0352°N, 23.0392°W) downwind of the Saharan desert. The fine-resolution features have been rendered as a correlative truth profile **x** consistent with the forward model F as indicated by equation (B1).

[52] **Acknowledgments.** This research was supported by the NOAA Joint Polar Satellite System Office (M. D. Goldberg) and the former Integrated Program Office (IPO), along with the STAR Satellite Meteorology and Climatology Division (F. Weng and I. Csiszar). AEROSE works in collaboration with the NOAA PIRATA Northeast Extension (PNE) project and is supported by the NOAA Educational Partnership Program grant NA17AE1625, NOAA grant NA17AE1623, JPSS and NOAA/NESDIS/STAR. We are grateful to several contributors to the S-NPP CrIMSS EDR validation effort, especially L. Zhou, M. Wilson, F. Iturbide-Sanchez, C. Tan, X. Xiong, H. Xie, J. Wei, F. Tilley, C. Brown, M. Petty (NOAA/NESDIS/STAR), and M. Feltz (UW/CIMSS). We also acknowledge the following collaborators for their contributions to the S-NPP validation data collection effort: B. Demoz and M. Oyola (Howard University, BCCSO and AEROSE); D. Wolfe (NOAA Earth System Research Laboratory, AEROSE); J. E. Wessel (Aerospace PMRF). We thank D. Holdridge and J. Mather and the U.S. DOE ARM Climate

Research Facility for its support of the satellite overpass radiosonde efforts. Data collection by The Aerospace Corporation was supported by NASA contract NNG11VH00B. We also thank two anonymous reviewers who provided many expert suggestions that we used to improve the overall quality of the paper. The views, opinions, and findings contained in this paper are those of the authors and should not be construed as an official NOAA or U.S. Government position, policy, or decision.

References

Adam, M., et al. (2010), Water vapor measurements by Howard University Raman Lidar during the WAVES 2006 campaign, *J. Atmos. Ocean. Tech.*, *27*(1), 42–60, doi:10.1175/2009JTECHA1331.1.

Anthes, R. A., et al. (2008), The COSMIC/FORMOSAT-3 mission: Early results, *Bull. Amer. Meteorol. Soc.*, *89*(3), 313–333, doi:10.1175/Bams-89-3-313.

Armstrong, B. H. (1968), Analysis of the Curtis-Godson approximation and radiation transmission through inhomogeneous atmospheres, *J. Atmos. Sci.*, *25*, 312–322.

Aumann, H. H., et al. (2003), AIRS/AMSU/HSB on the Aqua mission: Design, science objectives, data products, and processing systems, *IEEE Trans. Geosci. Remote Sens.*, *41*(2), 253–264.

Backus, G., and F. Gilbert (1970), Uniqueness in the inversion of inaccurate gross Earth data, *Philos. Trans. R. Soc. London, Ser. A*, *266*(1173), 123–192.

Barnet, C. (2009), NPOESS community collaborative calibration/validation plan for the NPOESS preparatory project CrIS/ATMS EDRs, *Tech. Rep. I30004, Ver. 1 Rev. B*, Integrated Program Office (IPO), Silver Spring, Maryland.

Cayla, F. R. (1993), IASI infrared interferometer for operations and research, in *High Spectral Resolution Infrared Remote Sensing for Earth's Weather and Climate Studies NATO ASI Series, I,* vol. 19, edited by A. Chedin, M. T. Chahine, and N. A. Scott, pp. 9–19, Springer-Verlag, Berlin Heidelberg.

Chahine, M. T., et al. (2006), AIRS: Improving weather forecasting and providing new data on greenhouse gases, *Bull. Am. Meteorol. Soc.*, *87*(7), 911–926.

Conrath, B. J. (1972), Vertical resolution of temperature profiles obtained from remote radiation measurements, *J. Atmos. Sci.*, *29*(7), 1262–1271.

Divakarla, M. G., C. D. Barnet, M. D. Goldberg, L. M. Mcmillin, E. Maddy, W. Wolf, L. Zhou, and X. Liu (2006), Validation of atmospheric infrared sounder temperature and water vapor retrievals with matched radiosonde measurements and forecasts, *J. Geophys. Res.*, *111*, D09S15, doi:10.1029/2005JD006116.

Divakarla, M. G. et al. (2014), The CrIMSS EDR Algorithm: Characterization, optimization and validation, *J. Geophys. Res. Atmos.*, doi:10.1002/2013JD020438.

Fetzer, E., et al. (2003), AIRS/AMSU/HSB validation, *IEEE Trans. Geosci. Remote Sens.*, *41*(2), 418–431.

Fetzer, E. J. (2006), Preface to special section: Validation of Atmospheric Infrared Sounder Observations, *J. Geophys. Res.*, *111*, D09S01, doi:10.1029/2005JD007020.

Gallery, W. O., F. X. Kneizys, and S. A. Clough, (1983), Air mass computer program for atmosperic transmittance/radiance calculation: FSCATM, *Environmental Research Papers 828*, Air Force Geophysics Laboratory, Hanscom AFB, AFGL-TR-83-0065.

Gambacorta, A., and C. Barnet (2013), Methodology and information content of the NOAA NESDIS operational channel selection for the Cross-Track Infrared Sounder (CrIS), *IEEE Trans. Geosci. Remote Sens.*, *51*, 3207–3216, doi:10.1109/TGRS.2012.2220369.

Goldberg, M. D., Y. Qu, L. M. Mcmillin, W. Wolf, L. Zhou, and M. Divakarla (2003), AIRS near-real-time products and algorithms in support of operational numerical weather prediction, *IEEE Trans. Geosci. Remote Sens.*, *41*, 379–389.

Hagan, D. E., and P. J. Minnett (2003), AIRS radiance validation over ocean from sea surface temperature measurements, *IEEE Trans. Geosci. Remote Sens.*, *41*(2), 432–441.

Hagan, D. E., C. R. Webster, C. B. Farmer, R. D. May, R. L. Herman, E. M. Weinstock, L. E. Christensen, L. R. Lait, and P. A. Newman (2004), Validating AIRS upper atmosphere water vapor retrievals using aircraft and balloon in situ measurements, *Geophys. Res. Lett.*, *31*, L21103, doi:10.1029/2004Gl020302.

Hilton, F., et al. (2012), Hyperspectral Earth observation from IASI: Five years of accomplishments, *Bull. Am. Meteorol. Soc.*, *93*(3), 347–370, doi:10.1175/BAMS-D-11-00027.1.

Hyland, R. W., and A. Wexler (1983), Formulations for the thermodynamic properties of the saturated phases of H_2O from 173.15 K To 473.15 K, *ASHRAE Trans.*, *2A*, 500–519.

Immler, F. J., J. Dykema, T. Gardiner, D. N. Whiteman, P. W. Thorne, and H. Vömel (2010), Reference quality upper-air measurements: Guidance for developing GRUAN data products, *Atmos. Meas. Tech.*, *3*, 1217–1231, doi:10.5194/amt-3-1217-2010.

Le Marshall, J., J. Jung, M. Goldberg, C. Barnet, W. Wolf, J. Derber, R. Treadon, and S. Lord (2008), Using cloudy AIRS fields of view in numerical weather prediction, *Aust. Met. Mag.*, *57*, 249–254.

Li, J., W. W. Wolf, W. P. Menzel, Z. Zhang, H. L. Huang, and T. H. Achtor (2000), Global soundings of the atmosphere from ATOVS measurements: The algorithm and validation, *J. Appl. Meteorol.*, *39*(8), 1248–1268.

Long, C. N., et al. (2013), ARM research in the equatorial western Pacific: A decade and counting, *Bull. Am. Meteorol. Soc.*, *94*(5), 695–708, doi:10:1175/BAMS-D-11-00137.1.

Lynch, R., J.-L. Moncet, and X. Liu (2009), Efficient nonlinear inversion for atmospheric sounding and other applications, *Appl. Optics*, *48*(10), 1790–1796.

Maddy, E. S., and C. D. Barnet (2008), Vertical resolution estimates in Version 5 Of AIRS operational retrievals, *IEEE Trans. Geosci. Remote Sens.*, *46*(8), 2375–2384.

Maddy, E. S., et al. (2011), Using MetOp-A AVHRR clear-sky measurements to cloud-clear MetOp-A IASI column radiances, *J. Atmos. Ocean. Tech.*, *28*, 1104–1116, doi:10.1175/JTECH-D-10-05045.1.

Maddy, E. S., et al. (2012), On the effect of dust aerosols on AIRS and IASI operational level 2 products, *Geophys. Res. Lett.*, *39*, L10809, doi:10.1029/2012GL052070.

Mango, S. A., R. E. Murphy, J. D. Cunningham, and P. Sablehouse, (2001), NPOESS NPP calibration and product validation plan, *IPO Technical Document Version 11*, 216 pp., Integrated Program Office.

Mather, J. H., and J. W. Voyles (2013), The ARM climate research facility: A review of structure and capabilities, *Bull. Amer. Meteorol. Soc.*, *94*(3), 377–392, doi:10.1175/BAMS-D-11-00218.1.

Miloshevich, L. M., H. Vömel, D. N. Whiteman, B. M. Lesht, F. J. Schmidlin, and F. Russo (2006), Absolute accuracy of water vapor measurements from six operational radiosonde types launched during AWEX-G and implications for AIRS validation, *J. Geophys. Res.*, *111*, D09S10, doi:10.1029/2005JD006083.

Minnett, P. J., R. O. Knuteson, F. A. Best, B. J. Osborne, J. A. Hanafin, and O. B. Brown (2001), The Marine-Atmospheric Emitted Radiance Interferometer (M-AERI): A high-accuracy, sea-going infrared spectroradiometer, *J. Atmos. Oceanic Technol.*, *18*, 994–1013.

Mollner, A. K., J. E. Wessel, K. M. Gaab, D. M. Cardoza, S. D. Lalumondiere, and D. L. Caponi, (2013), Ground truth data collection for assessment of ATMS/CrIS Sensors Aboard Suomi-NPP, *Aerospace Report ATR-2013(5758)-2*, The Aerospace Corporation, Electronics And Photonics Laboratory, prepared for NASA/GSFC, Contract No. Nng11Vh00B.

Moncet, J.-L., et al. (2005), *Joint Polar Satellite System (JPSS) Algorithm Theoretical Basis Document for the Cross Track Infrared Sounder (CrIS) Volume II, Environmental Data Records (EDR)*, Joint Polar Satellite System (JPSS), Greenbelt, Maryland, code 474-00056, http://npp.gsfc.nasa.gov/science/sciencedocuments/2013-01/474-00056_RevABaseline.pdf.

Morris, V., P. Clemente-Colón, N. R. Nalli, E. Joseph, R. A. Armstrong, Y. Detrés, M. D. Goldberg, P. J. Minnett, and R. Lumpkin (2006), Measuring trans-Atlantic aerosol transport from Africa, *Eos Trans. Agu*, *87*(50), 565–571.

Nalli, N. R., et al. (2006), Ship-based measurements for infrared sensor validation during Aerosol and Ocean Science Expedition 2004, *J. Geophys. Res.*, *111*, D09S04, doi:10.1029/2005JD006385.

Nalli, N. R., P. J. Minnett, E. Maddy, W. W. Mcmillan, and M. D. Goldberg (2008), Emissivity and reflection model for calculating unpolarized isotropic water surface leaving radiance in the infrared. 2: Validation using Fourier transform spectrometers, *Appl. Optics*, *47*(25), 4649–4671.

Nalli, N. R., et al. (2011), Multiyear observations of the tropical Atlantic atmosphere: Multidisciplinary applications of the NOAA Aerosols and Ocean Science Expeditions (AEROSE), *Bull. Am. Meteorol. Soc.*, *92*, 765–789, doi:10.1175/2011BAMS2997.1.

Nalli, N. R., C. D. Barnet, A. Gambacorta, E. S. Maddy, H. Xie, T. S. King, E. Joseph, and V. R. Morris (2013), On the angular effect of residual clouds and aerosols in clear-sky infrared window radiance observations. 2. Satellite experimental analyses, *J. Geophys. Res. Atmos.*, *118*, 1–16, doi:10.1029/2012JD018260.

Newman, S. M., et al. (2012), The Joint Airborne IASI Validation Experiment: An evaluation of instrument and algorithms, *J. Quant. Spectrosc. Radiat. Transfer*, *113*, 1372–1390.

Pagano, T. (2013), AIRS Project Status, Sounder Science Team Meeting, NASA/JPL, Pasadena, Ca, http://airs.jpl.nasa.gov/Documents/Science_Team_Meeting_Archive/2013_05/Slides/Pagano_Airs_Proj_Status_May_2013.Pdf.

Pougatchev, N. (2008), Validation of atmospheric sounders by correlative measurements, *Appl. Optics*, *47*(26), 4739–4748.

Reale, A., F. Tilley, M. Ferguson, and A. Allegrino (2008a), NOAA operational sounding products for advanced TOVS, *Int. J. Remote Sens.*, *29*(16), 4615–4651, doi:10.1080/01431160802020502.

Reale, O., J. Susskind, R. Rosenberg, E. Brin, E. Liu, L. P. Riishojgaard, J. Terry, and J. C. Jusem (2008b), Improving forecast skill by assimilation of quality-controlled AIRS temperature retrievals under partially cloudy conditions, *Geophys. Res. Lett.*, *35*, L08809, doi:10.1029/2007GL033002.

Reale, O., W. K. Lau, J. Susskind, E. Brin, E. Liu, L. P. Riishojgaard, M. Fuentes, and R. Rosenberg (2009), AIRS impact on the analysis and forecast track of tropical cyclone Nargis in a global assimilation and forecast system, *Geophys. Res. Lett.*, *36*, L06812, doi:10.1029/2008GL037122.

Reale, O., K. M. Lau, J. Susskind, and R. Rosenberg (2012a), AIRS impact on analysis and forecast of an extreme rainfall event (Indus River Valley, Pakistan, 2010) with a global data assimilation and forecast system, *J. Geophys. Res.*, *117*, D08103, doi:10.1029/2011JD017093.

Reale, T., B. Sun, F. H. Tilley, and M. Pettey (2012b), The NOAA Products Validation System (NPROVS), *J. Atmos. Ocean. Tech.*, *29*, 629–645, doi:10.1175/JTECH-D-11-00072.1.

Rodgers, C. D. (1976), Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation, *Rev. Geophys. Space Phys.*, *14*(4), 609–624.

Rodgers, C. D. (1990), Characterization and error analysis of profiles retrieved from remote sounding measurements, *J. Geophys. Res.*, *95*(D5), 5587–5595.

Rodgers, C. D., and B. J. Connor (2003), Intercomparison of remote sounding instruments, *J. Geophys. Res.*, *108*(D3), 4116, doi:10.1029/2002JD002299.

Seidel, D. J., et al. (2009), Reference upper-air observations for climate: Rationale, progress, and plans, *Bull. Amer. Meteorol. Soc.*, *90*(3), 361–369, doi:10.1175/2008BAMS2540.1..

Seidel, D. J., N. P. Gillett, J. R. Lanzante, K. P. Shine, and P. W. Thorne (2011), Stratospheric temperature trends: Our evolving understanding wires, *Clim. Change*, *2*, 592–616, doi:10.1002/wcc.125.

Smith, W. L., et al. (1996), Observations of the infrared radiative properties of the ocean: Implications for the measurement of sea surface temperature via satellite remote sensing, *Bull. Am. Meteorol. Soc.*, *77*, 41–51.

Smith, W. L., H. Revercomb, G. Bingham, A. Larar, H. Huang, D. Zhou, J. Li, X. Liu, and S. Kireev (2009), Technical note: Evolution, current capabilities, and future advance in satellite nadir viewing ultraspectral IR sounding of the lower atmosphere, *Atmos. Chem. Phys*, *9*, 5563–5574.

Stephens, G. L. (1994), *Remote Sensing of the Lower Atmosphere: An Introduction*, 523 pp., Oxford Univ. Press, New York.

Strow, L. L., S. E. Hannon, S. D. Souza-Machado, H. E. Motteler, and D. Tobin (2003), An overview of the AIRS radiative transfer model, *IEEE Trans. Geosci. Remote Sens.*, *41*(2), 303–313.

Stull, R. B. (2000), *Meteorology for Scientists and Engineers,* 2nd ed., Brooks/Cole, Cengage Learning, Belmont, Calif.

Sun, B., A. Reale, D. J. Seidel, and D. C. Hunt (2010), Comparing radiosonde and COSMIC atmospheric profile data to quantify differences among radiosonde types and the effects of imperfect collocation on comparison statistics, *J. Geophys. Res.*, *115*, D23104, doi:10.1029/2010JD014457.

Sun, B., A. Reale, S. Schroeder, D. J. Seidel, and B. Ballish (2013), Toward improved corrections for radiation-induced biases in radiosonde temperature observations, *J. Geophys. Res. Atmos.*, *118*, 4231–4243, doi:10.1002/jgrd.50369.

Susskind, J., J. Rosenfield, and D. Reuter (1983), An accurate radiative transfer model for use in the direct physical inversion of HIRS2 and MSU temperature sounding data, *J. Geophys. Res.*, *88*(C13), 8550–8568.

Susskind, J., C. D. Barnet, and J. M. Blaisdell (2003), Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds, *IEEE Trans. Geosci. Remote Sens.*, *41*(2), 390–409.

Susskind, J., J. Blaisdell, L. Iredell, and F. Keita (2011), Improved temperature sounding and quality control methodology using AIRS/AMSU data: The AIRS science team version 5 retrieval algorithm, *Ieee Trans. Geosci. Remote Sens.*, *49*(3), 883–907, doi:10.1109/TGRS.2010.2070508.

Taylor, J. P., et al. (2008), EAQUATE: An international experiment for hyperspectral atmospheric sounding validation, *Bull. Amer. Meteorol. Soc.*, *89*(2), 203–218, doi:10.1175/BAMS-89-2-203.

Tobin, D. C., H. E. Revercomb, R. O. Knuteson, B. M. Lesht, L. L. Strow, S. E. Hannon, W. F. Feltz, L. A. Moy, E. J. Fetzer, and T. S. Cress (2006a), Atmospheric Radiation Measurement site atmospheric state best estimates for atmospheric infrared sounder temperature and water vapor retrieval validation, *J. Geophys. Res.*, *111*, D09S14, doi:10.1029/2005JD006103.

Tobin, D. C., et al. (2006b), Radiometric and spectral validation of Atmospheric Infrared Sounder observations with the aircraft-based scanning high-resolution interferometer sounder, *J. Geophys. Res.*, *111*, D09S02, doi:10.1029/2005JD006094.

Turtiainen, H., H. Jauhiainen, J. Lentonen, P. Survo, V.-P. Viitanen, and W. F. Dabberdt (2010), Upper atmosphere humidity measurements with the APS sensor: 1st progress report on the Vaisala Reference Radiosonde Program, paper presented at 15th Symposium on Meteorological Observation and Instrumentation, 90th Annual Meeting, Am. Meteorol. Soc., Atlanta, GA, https://ams.confex.com/ams/90annual/techprogram/paper_161720.htm.

Wallace, J. M., and P. V. Hobbs (1977), *Atmospheric Science: An Introductory Survey*, Academic Press, San Diego, Calif.

Wang, J., T. Hock, S. A. Cohn, C. Martin, N. Potts, T. Reale, B. Sun, and F. Tilley (2013), Unprecedented upper-air dropsonde observations over Antarctica from the 2010 Concordiasi Experiment: Validation of satellite-retrieved temperature profiles, *Geophys. Res. Lett.*, *40*, 1231–1236, doi:10.1002/grl50246.

Weng, F., X. Zou, X. Wang, S. Yang, and M. D. Goldberg (2012), Introduction to Suomi National Polar-Orbiting Partnership Advanced Technology Microwave Sounder for numerical weather prediction and tropical cyclone applications, *J. Geophys. Res.*, *117*, D19112, doi:10.1029/2012JD018144.

Wessel, J., R. W. Farley, A. Fote, Y. Hong, G. A. Poe, S. D. Swadley, B. Thomas, and D. J. Boucher (2008), Calibration and validation of DMSP SSMIS lower atmospheric sounding channels, *IEEE Trans. Geosci. Remote Sens.*, *46*(4), 946–961, doi:10.1109/tgrs.2008.917132.

Wexler, A. (1976), Vapor pressure formulation for water in range 0 to 100°C: A Revision, *Res. Natl. Bur. Stand. U.S., Sect. A*, *80*, 775–785.

Whiteman, D. N., et al. (2006), Analysis of Raman lidar and radiosonde measurements from the AWEX-G field campaign and its relation to Aqua validation, *J. Geophys. Res.*, *111*, D09S09, doi:10.1029/2005Jd006429.

Xie, H., N. R. Nalli, S. Sampson, W. W. Wolf, J. Li, T. J. Schmit, C. D. Barnet, E. Joseph, V. R. Morris, and F. Yang (2013), Integration and ocean-based prelaunch validation of GOES-R Advanced Baseline Imager legacy atmospheric products, *J. Atmos. Ocean. Tech.*, *1743–1756*(8), doi:10.1175/JTECH-D-12-00120.1.

Zhou, Y. P., K. M. Lau, O. Reale, and R. Rosenberg (2010), AIRS impact on precipitation analysis and forecast of tropical cyclones in a global data assimilation and forecast system, *Geophys. Res. Lett.*, *37*, L02806, doi:10.1029/2009Gl041494.